FUTURE PRICE INDEX OF FARM PRODUCTS BASED ON CLIMATE FACTORS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

NURŞEN KARASU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
FINANCIAL MATHEMATICS

JULY 2024

Approval of the thesis:

FUTURE PRICE INDEX OF FARM PRODUCTS BASED ON CLIMATE FACTORS

submitted by **NURŞEN KARASU** in partial fulfillment of the requirements for the degree of **Master of Science in Financial Mathematics Department, Middle East Technical University** by,

Prof. Dr. A.Sevtap Selçuk Kestel
Dean, Graduate School of **Applied Mathematics** __________

Prof. Dr. A.Sevtap Selçuk Kestel
Head of Department, **Financial Mathematics** __________

Prof. Dr. Erkan Erdil
Supervisor, **Economics, METU** __________

Prof. Dr. Kasırga Yıldırak
Co-supervisor, **Actuarial Sciences, Hacettepe University** __________

Examining Committee Members:

Prof. Dr. A.Sevtap Selçuk Kestel
Actuarial Sciences, METU __________

Prof. Dr. Erkan Erdil
Economics, METU __________

Prof. Dr. Kasırga Yıldırak
Actuarial Sciences, Hacettepe University __________

Assist. Prof. Dr. Büşra Zeynep Temoçin
Actuarial Sciences, METU __________

Assist. Prof. Dr. Başak Bulut Karageyik
Actuarial Sciences, Hacettepe University __________

Date: __________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: NURŞEN KARASU

Signature :

ABSTRACT

FUTURE PRICE INDEX OF FARM PRODUCTS BASED ON CLIMATE FACTORS

Karasu, Nurşen
M.S., Department of Financial Mathematics
Supervisor : Prof. Dr. Erkan Erdil
Co-Supervisor : Prof. Dr. Kasırga Yıldırak

July 2024, 101 pages

Climate conditions have a big impact on the yield of farm products, hence the prices. This thesis makes price prediction of majorly traded grains, Wheat, Barley and Corn, based on major climate conditions, total precipitation and dew point, levels of which are taken from European Centre for Medium-Range Weather Forecast (ECMWF) database of Konya, Polatli, Yozgat, Adana and Urfa where major production is held, by using a smart, machine learning method of ensemble training and compares with the price prediction results retrieved by simple regression method. These predictions would help future prices in futures commodity markets to be determined and any desired future price index can then be retrieved through these settled end of day future prices, by applying weights decided by index structurers. The study also confirms that ensemble training method could be used for future price prediction even when the statistic significance of data is low, therefore a successful simple regression optimization methodology is hard to apply. On the other hand, high volatile inflation rate and exchange rate would lead the predictions to deviate outside the accepted limits so the model studied in this paper is believed to be a better fit in stabilized economies. This thesis study offers important clues for producers for their production preferences, policy makers to build production planning, insurance companies to make climate risk mappings and financial instrument traders to make reasonable pricing so to prevent volatility in commodity market. Finally, all thesis study outputs are

expected to serve the purpose of maintaining a sustainable agricultural production.

Keywords: Climate Index, Machine Learning Models, Ensemble Training Methods, Future Price Index on Farm Products, Sustainability, Agricultural Risk Index

ÖZ

İKLİM FAKTÖRLERİNE DAYALI TARIM ÜRÜNLERİ VADELİ FİYAT ENDEKSİ

Karasu, Nurşen
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi : Prof. Dr. Erkan Erdil
Ortak Tez Yöneticisi : Prof. Dr. Kasırga Yıldırak

Temmuz 2024, 101 sayfa

İklim koşullarının tarım ürün hasılatına, dolayısıyla tarım ürünü fiyatlarına büyük bir etkisi bulunmaktadır. Bu tez, en çok alım-satım yapılan ürünler olan Buğday, Arpa ve Mısır üzerine, en yüksek üretimin yapıldığı yerler olan Konya, Yozgat, Polatli, Adana ve Urfa'daki en önemli iklim koşulları sayılabilecek, European Center for Medium-Range Weather Forecast'ten (ECMWF) alınan toplam yağış ve çiğ oluşma derecesi seviyelerinin etkisini, akıllı, makine öğrenmesi yöntemi olan topluluk öğrenmesini kullanarak fiyat tahminlemesinde bulunmakta ve basit regresyon modeli fiyat tahmin sonuçları ile karşılaştırmaktadır. Bu tahminler vadeli emtia piyasalarındaki vadeli fiyatların belirlenmesine yardımcı olacak ve ardından uzlaşılan gün sonu vadeli fiyatlar üzerinden, endeks oluşturanlar tarafından karar verilecek ağırlıklar uygulanarak istenilen vadeli fiyat endeksi kurulabilecektir. Bu çalışma ayrıca, istatistiksel anlamlılığı düşük olan veri setinde yani başarılı bir basit regresyonun metodolojisinin uygulanmasının zor olduğu durumlarda dahi topluluk öğrenmesi yönteminin vadeli fiyat tahminlemesinde kullanılabileceğini teyit etmektedir. Diğer taraftan, enflasyon ve kur dalgalanmalarının yüksek olduğu durumlarda tahminler kabul edilebilir sınırların üzerinde sapmalar göstereceğinden bu belgede çalışılan modelin stabil ekonomiler için daha uygun olduğuna inanılmaktadır. Bu tez, üreticilere üretim tercihleri yönünden, politika yapıcılara üretim planlaması, sigorta şirketlerine risk haritalandırmaları yönünden ve finansal enstrüman alım satımcılarına makul fiyatlandırma yapmaları

ve böylelikle emtia piyasasındaki dalgalanmanın önüne geçilmesi yönünden önemli ipuçları sunmaktadır. Nihayetinde tüm tez çıktılarının sürdürülebilir tarım üretiminin korunması amacına hizmet etmesi beklenmektedir.

Anahtar Kelimeler: İklim Endeksi, Makine Öğrenmesi Modelleri, Topluluk Öğrenmesi Yöntemleri, Tarım Ürünleri Vadeli Fiyat Endeksi, Sürdürebilirlik, Tarımsal Risk Endeksi

To my grandmother,
who always worried about me when studying and wondered
if I would end up being a professor (or what?)

ACKNOWLEDGMENTS

I would like to express my gratitudes to my thesis supervisors, distinguished not only for their distinct reputations but also for their empathetic, understanding and supporting guidance. Their encouragement and valuable advices were my lighthouse during the development and preparation of this thesis. Prof. Dr. Erkan Erdil was very kind to accept my request on being my advisor, despite his busy schedule on his own department and still could respond to my numerous questions in shortest time. Prof. Dr. Kasırga Yıldırak was very sincere on his effort and he brightened my path with his innovative ideas so I could be able to launch an uncommon approach. Dr. Asım Özalp from the Actuarial Science department at Hacettepe University offered continuous support and patience, sharing invaluable insights and assistance throughout my study. Additionally, I would like to express my heartfelt gratitude to Prof. Dr. Devin Sezer for motivating me to persevere during taking his course in this programme.

I am also indebted to my midlife crisis, which steered me towards the academic world instead of a career in banking, allowing me the opportunity to contribute, albeit modestly, to the scholarly literature. My husband Nazım has my deepest gratitude for his support on this decision of mine in the beginning and during moments of doubt while completing this program especially after giving birth to our dear son Gediz. I feel so lucky to have them in my life.

And I thank myself, for not giving up on hopes, for not being intimidated by numerous unknowns and for my never-can-be-tamed romantism... all inherited from my family, together with my maiden name ALTUNCI, teaching me morals and ethics so I could endure against all challanges by keeping my self-respect and dignity, hopefully to be able to serve science.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AEPDI	Agricultural Economics and Policy Development Institute
APW	Australian Premium White
ASW	Australian Standard White
CBOT	Chicago Board of Trade
CBT	Central Bank of the Republic of Türkiye
CME	Chicago Mercantile Exchange
CPT	Carriage Paid to
CWRS	Canada Western Red Spring
ÇEY	Çiftçi Eğitimi ve Yayım Şube Müdürlüğü
ÇKS	Çiftçi Kayıt Sistemi - Farmer Registration System
EWR	Electronic Warehouse Receipt
ECMWF	European Centre for Medium-Range Weather Forecast
EU	European Union
FAO	Food and Agriculture Organization of United Nations
FED	Federal Reserve
FOB	Free on Board
HRW	Hard Red Winter
NMEX	National Commodity and Derivatives Exchanges
SRW	Soft Red Winter
TAGEM	Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğü - General Directorate of Agricultural Research and Policies
TARSIM	Tarım Sigortaları Havuz İşletmesi A.Ş. - Agricultural Insurance Pool Management Company
TEPGE	Tarımsal Ekonomi ve Politika Geliştirme Enstitüsü Müdürlüğü - Agricultural Economic and Policy Development Institute
TMEX	Turkish Mercantile Exchange
TMO	Toprak Mahsulleri Ofisi - Turkish Grain Board
TURKSTAT	Turkish Statistical Institute
USDA	United States Department of Agriculture

CHAPTER 1

INTRODUCTION

Traditionally, the price of farm products, like many other goods and services, depends on their availability in the market and the urgency of their need. On the other hand, there are other factors affecting availability, such as storage conditions that extend the expiration date and effective production methods that help achieve multiple outputs from one unit of input. In addition to current inflation on food and future demand, climate factors play a determining role in farm product pricing.

In today's world, however, the demand and supply of farm products meet not only in physical markets but also on organized electronic platforms. As farm products started to be traded in organized platforms and as investors improved in terms of sophistication and number, the need for fair and transparent information gathering methods became important and the need for a prediction model based on a single factor has arisen.

By proposing a price modeling framework and applying the resulting model on mercantile exchange indices, the proposed study seeks to develop a relationship between climate factors and the availability of farm products to realize future price estimations. Since spot prices of farm products have already been determined under market conditions, this research aims to create a future price index for farm products by making an accurate forecast of future prices for specific farm products in Turkey, taking climate effects into consideration, such as weather conditions, temperature, and rain and/or snowfall. Establishing a well-running model of climate factors and farm product prices during a particular time frame is the focus of this analysis.

1.1 Statement of the Problem

"Climate factors are important determinants of farm product prices, which can be measured by establishing the relationship via machine learning systems," is the main statement of this study. Climate factors are important on the production side not only because they determine the harvested amount depending on weather conditions such as floods, frosts, and droughts but also affect the storage conditions and the precautions required, such as agricultural insurance, making the input cost higher. This can cause the incentives of entrepreneurs to cultivate land to fluctuate; hence, even the number of suppliers depends on climate conditions.

1.2 Objectives and Research Question

The central purpose of this study is to establish the relationship of farm product prices and climate conditions to develop a future price index of predetermined farm products. In doing so, the study targets to accomplish the following sub-objectives:

- Establishing the reference prices to increase transparency in the agricultural commodity markets so improving the confidence in the trading process.
- Having an almost accurate estimation on future product prices to help the policy makers to understand the need for public interventions and subventions.
- Presenting an indicator for insurance companies while creating their insurance risk models.
- Stating a reference for real sector entrepreneurs while giving future investment decisions, so the industry can pace accordingly.
- Building a new price index as a possible underlying asset for index based derivative products so a new market instrument can be created.

In addition, this research is planned to be handled upon multiple layers as such:

- Creating a time series model using historical data on both climate variables such as temperatures, rain and snow falls versus farm product prices.

- Stating a correlation between the climate factor measures and the production levels to bring an understanding about the changing product price levels.
- By using this correlation, reaching an accurate estimation of future prices on specific farm products, with the given data of weather forecasts.
- Producing a future price index based on these almost accurate forecasts depending on climate conditions estimations.

1.3 Data Analysis

Since all climate factors are measurable in terms of numbers, the research is planned to be based on quantitative data. Climate factors such as weather temperature, dew point (which indicates soil humidity), snowfall, and rainfall amount (in terms of precipitation) will be analyzed. Additionally, production numbers and prices for Turkey will be extracted for the set time interval. The relationship between climate conditions and production numbers, hence the price affected by this supply will be modeled. By deriving such a model, the aim is to use the correlation between climate conditions and prices to make an accurate estimation of future prices for specific farm products. With these accurate estimations, a future price index could be derived.

The detailed information on data set is given in Section 3.1

1.4 Methodology

To minimize the basis risk associated with index derivatives is the most important criteria to evaluate the performance of cross hedging strategies. That is why it is intended to set minimum attainable value of basis risk to choose the weights assigned to the variables constructing the index itself. There are various observations that could be employed. Most of them is the model data obtained from ECMWF interim mesoscale dataset such as relative humidity (dew point), total precipitation and observed yield-price data.

On this basis, the research formulated mainly in three steps;

- **First - Literature review**: To be able to have a theoretical understanding on the issue, literature review has been done. In literature review part, the aim is to understand the historical background of the farm products in terms of production and harvesting methods and their availabilities, government interventions, pricing motivations behind, global tendency towards consumption of those products and the structural changes of their availability and demand. Both local and global data is analyzed. Preview studies on similar field is examined.

- **Second - Preliminary data collection**: To be able put a general information, preliminary data is collected from national or international web-based sources like international exchanges, databanks, state agencies, statistical information, reports and other information collected from related national web-based sources. Data sources are given in Table A.1

- **Third - Time-series model**: Constrained optimization problem with basis risk being the objective to minimize. MATLAB is the primary software

By following these steps, design of the dynamic optimization problem and its solution is targeted to be developed throughout the thesis work. Details are explained in Section 3.2.

1.5 Background in Literature

To the best of our knowledge, this thesis is the first to predict future prices of three grains using two weather condition variables from ECMWF, employing the ensemble tree learning method.

The elements that are making the model studied in this thesis unique can be counted as follows:

- The data used is ECMWF data.
- There are more than one province.
- There are more than one type of weather information.
- There are more than one type of farm product.

- This study demonstrates the potential to utilize climate factors for future price predictions, even when the time series data efficiency is lower than desired.
- The model examined in this study leverages climate conditions to predict future prices using an advanced machine learning technique known as the ensemble tree learning method.

1.6 Content

This thesis is organized into four chapters, including an introductory section presenting the core of the thesis by outlining the aim, background, and methodology in Chapter 1. In Chapter 2, general information on the selected farm products is given, including their areas of usage, the physical conditions required for production, and the phenological atlas in Turkey showing the climate requirements per period. It also includes other important data to understand the significance of these products, such as the planting ratio, historical importance, public initiatives and insurance applications. Chapter 3 is based on the model structured to link climate conditions to yield and then yield to price, encompassing two phases. It details the methodology, data structure, interpretation of statistics and results. Chapter 4 offers a general interpretation by comparing the results, discussing the policy implications, study highlights and limitations and explaining the outcomes, ultimately leading to the conclusion.

CHAPTER 2

GENERAL INFORMATION ON THE FARM PRODUCTS

To comprehend the climate effects on farm product prices, it is essential to grasp the entire cultivation process of crops, starting from sowing the seeds through to harvesting.

Since the beginning of agriculture on earth, the major concern has been growing crops and storing them so to maintain life on earth [1]. As stated in the book "Wheat Improvement, Food Security in a Changing Climate" edited by Matthew P. Reynolds and Hans-Joachim Braun in 2022, agriculture started to take place by domestication of wild plants at least 12,000 years ago in the Neolithic age. This resulted in 100 species that is cultivated now; in addition to that 7000 of plant species such as medical plants, herbs and spices are considered semi-cultivated. [36].

The conditions involve not only temperatures of the weather but also the humidity of the soil, the altitude and the slope of the land. Among these conditions, the meteorological data such as weather temperature, humidity, rain/snowfall, high wind and when applicable, sudden meteorological shocks needs to be taken into account during all processes of production. In this study, the two pivotal factos of precipitation and dew point is taken into consideration as they are measuable and recordable figures available.

Before analyzing the data, it would be better to understand the conditions needed for an efficient harvest. In this chapter the product characteristics and their dependence on the weather, soil conditions and timings are analyzed in different aspects. The dependence of humans and animals to those crops are also understood. Their importance of their production on human nutrition as well as their source of living are also highlighted. After the historical background of those crops and their crucial place among life, the supporting public policies and their insurance application in Turkey is explained as to understand the production trend, phenological effects are needed to be considered so to build the periods of time series.

This chapter mainly provides an insight on the selected farm products on these aspects.

2.1 Wheat

Wheat is a fundamental food crop in human nutrition, providing a rich source of carbohydrates and representing a significant portion of crop planting, akin to corn and rice. It holds the distinction of being the most widely cultivated plant in Turkey and globally, owing not only to its nutritional efficacy but also its substantial contribution to rural income [2]. While wheat production is primarily concentrated in Europe followed by Asia, its pivotal role in meeting the demands of a rapidly growing population persists, despite a declining production trend [27]. Therefore, it is essential to comprehend key characteristics of wheat, including its structure, predominant uses, and meteorological requirements for optimal growth.

2.1.1 Physical Characteristics of Wheat

Since its domestication approximately 10,000 years ago, wheat has been vital to global food security. Today, it provides a fifth of the world's food calories and protein. As the most extensively grown crop worldwide, wheat is cultivated on 217 million hectares each year [27]. According to USDA, this number is estimated to be 222 million ha for 2021/2022 season [32]. As per the studies of Ministry of Agriculture of Turkey [4], wheat is produced by around 2.9 millions of enterprises and stands as the earning of living of around 15 millions of people in Turkey [22], the numbers are growing as the population increase.

The cultivation of wheat is taking place under dry conditions so that the yield is low and the income of the wheat producers is comparatively low and in some areas wheat production cannot be substituted by another alternative [4].

Meeting the physical requirements of wheat cultivation is relatively straightforward, evident from its widespread cultivation worldwide. In Turkey, the Ministry of Agriculture has been instrumental in providing informative resources to support local farmers, detailing the characteristics of various wheat types. For instance, in publications like the handbook on wheat cultivation by the Agricultural Department of Bursa Governorship, it is highlighted that wheat is not overly selective in its soil and climate requirements, making it adaptable to nearly all types of land. This adaptability is a key factor contributing to its extensive cultivation globally. Ideally, wheat thrives in deep clayish and loamy soils, rich in nutrients, well-ventilated, and with proper drainage and a neutral pH level [2].

Parallel to its non-selective requirements and straightforward characteristics, wheat cultivation occurs globally, with countries planting the crop sequentially according to their respective seasons. However, wheat is susceptible to various biological threats, including physiological and biochemical factors, which can have a significant impact on crop yield and quality. These threats, such as pathological and entomological diseases, underscore the importance for farmers to ensure proper mineral content in crops to facilitate grain development from germination through to harvest [54].

Considering the possible link of the exposed biochemical factors to the weather conditions, the need to collect the meteorological data and the cultivation outcomes is here again important.

2.1.2 Areas of Usage of Wheat

With its appropriate nutritional value and convenience for maintaining and processing, wheat stands as the basic nutrient in around 50 countries and provides approximately 20% of the total calories received from vegetative foods for the world population. This ratio is 53% for our country as per the studies of Ministry of Agriculture of Turkey in 2016[4]. 78.1% of Turkey's wheat consumption is for food, 12.4% for feed and 6.6% for seeds according to TEPGE's report in 2022 [32].

Wheat is classified under 2 subgroups according to their convenience to produce bread or pasta, respectively named as "milling wheat" and "durum wheat". The 51% of the cultivated land is wheat and durum wheat is the 19% of the wheat planted land as declared by Ministry of Agriculture in 2012 [22]. In 2021/2022 season, the cultivated land is estimated to be 67 million da in Turkey for wheat where 17.8% of it is allocated for durum wheat [33].

Wheat grains are used to produce flour, couscous, pasta and starch in terms of human nutrition. Wheat stalks on the the other hand are used in animal feed industry as well as the paper and cardboard industry [2].

The content of wheat in various important wheat markets around the world, categorized by type and area of usage, is presented in Table 2.1). The protein type is a crucial determinant in understanding the optimal utilization of wheat. This table, derived from the S&P Global "Global Grains and Oilseeds" study [39], last updated in January 2024, highlights slight differences among countries and how these variances can define distinct types of wheat. High protein content in wheat makes it suitable for pasta production, whereas wheat with low protein content is considered lower quality and is typically used for breadmaking when blended with other protein-rich wheat varieties.

Table 2.1: The Protein Content of Some Wheat Types and Their Areas of Usage

Type & Classification	Region	Protein %	Areas of usage
ASW (Australian Standard White) Wheat FOB Australia	Adelaide, North Australia, Australia	Min 9.5%	Flatbread and cookie/biscuits.
APW (Australian Premium White) Wheat FOB Australia	Adelaide, North Australia, Australia	10.5%	Bread or flour for general purpose.
CWRS (Canada Western Red Spring) Wheat FOB	Quebec, Kanada	13.5%	The nutrient quality of the low protein contained wheat is enhanced by getting mixed with this wheat and altogether used for pasta and bread.
EU milling wheat CPT France basis Rouen	Rouen, Haute-Normandie, France	11%	Bread or flour for general purpose. It is an A quality wheat
EU milling wheat FOB Hamburg	South and West Germany	11.5%	Bread or flour for general purpose
EU milling wheat FOB Constanta	South and Southeast Romania	12.5%	Bread or flour for general purpose
FOB Black Sea wheat (Ukraine)	Odessa Oblastı, Ukraine	11%	Bread or flour for general purpose
FOB Black Sea wheat (Russia)	Krasnodar Krayı and Rostov Oblastı, Russia	12.5%	Bread or flour for general purpose

2.1.3 Required Weather Conditions for Production for Wheat

During its initial vegetative stage, which includes germination and tillering, wheat thrives in temperatures ranging from 5 to 10 degrees Celsius and requires a relative humidity above 60%. High temperatures and excessive light are unnecessary during this period. As the crop progresses into the generative stage, characterized by bolting, temperatures of 10 to 15 degrees Celsius and a relative humidity of 65% significantly promote its development, as outlined in the Agricultural Department Handbook of the Bursa Governorship [2]. Before earing, wheat necessitates ample light and high humidity. Following fertilization (full earring), it benefits from low humidity and high temperatures to ensure a qualitative and abundant harvest.

The handbook further notes that when wheat has 3 to 5 leaves and 1 to 2 tillers just before winter, it can withstand temperatures as low as -35 degrees Celsius. This is crucial information for determining the optimal sowing time, as wheat lacking 3 to 5 leaves may be susceptible to damage from cold and harsh winter conditions.

Overwatering during the vegetative period can result in lodging, a problem also observed with heavy irrigation during the late yield formation period. According to the FAO, for high yields, water requirements, measured as evapotranspiration (ETm), range from 450 to 650 mm, depending on climate and the length of the growing period. The crop coefficient (kc), which relates maximum evapotranspiration (ETm) to reference evapotranspiration (ETo), varies throughout different growth stages [16]. Wheat utilizes 50 to 60% of soil water until the next irrigation, with depletion occurring more rapidly during the ripening period. Optimal root development occurs when the root zone is saturated with water throughout the sowing period.

FAO also highlights the importance of avoiding water deficits during the flowering period and the month or two following sowing, particularly for spring wheat. However, water has less impact on yields during the late period when the crop is ripening, as the remaining soil water is typically sufficient [16].

Regarding weather conditions for agriculture in Turkey, as per the Ministry of Agriculture document from 2016, the harvest occurs over a 3.5-month period between mid-May and mid-August, varying by region. According to the same document, which serves as an educational resource for farmers, the optimal time to commence cultivation is when the humidity per grain reaches 13.5%. Similarly, wheat storage should not exceed a humidity of 13% [4]. Harvesting should commence when the plant has turned completely yellow, and the grain has hardened.

Additionally, it's noted that milling wheat demonstrates greater resilience to harsh winter conditions compared to durum wheat, although the resistance capabilities can vary within milling wheat varieties. The main necessities can be summarized as shown in Table 2.2 [10]

Table 2.2: Necessary Conditions for Wheat Production

Wheat Type	**Rainfall**	**Avg Temp/Humidity**	**Altitude**
Wheat for Bread	Annual: 30 mm<X<1600mm	Effective temperature threshold [1]= 5 C for min 1000 days Avg Temp Jan: >-10 Celcius Avg Temp Feb: >-10 Celcius Avg Temp Sept: >8 Celcius	<1600 m
Wheat for Pasta	May and June: 15 mm<X<100mm	Avg Temp May: >17 Celcius Avg Humidity May: <55% Avg Humidity Jun: <66%	<1450mm
Wheat for Biscuits	April, May and June : >145 mm		

2.1.4 Significance of Wheat

Wheat is undeniably one of the most widely sown crops globally, with production reaching 773 million kilograms in 2022. Nearly half of this production, 48.8%, is concentrated in China, the EU, India, Russia, and the USA. Leading the world in wheat exports are Russia, the EU, Australia, the USA, Ukraine, and Argentina, respectively [32].

A study conducted by the Ministry of Agriculture in June 2022, focusing on Agricultural Products Markets, revealed that the highest wheat productions occurred in Konya (9%), Şanlıurfa (6.7%), and Tekirdağ (5.8%) in 2021 [32]. Moreover, in terms of wheat planting land, Konya claimed the top spot with a 9.8% share, followed by Şanlıurfa with 5.1%, and Ankara with 4.5% during the 2021/2022 season. Other significant wheat-producing provinces include Adana, Edirne, Ankara, Mardin, Diyarbakır, Kahramanmaraş, and Kırklareli [32]. Among these provinces, Konya, Şanlıurfa, Adana, and Ankara are selected for this thesis study based on their total production rates during the specified time interval.

As of the 2022/23 harvesting period, Turkey's wheat planting area accounts for 3.1%

[1]The sum of temperature differences above specified temperature during the days experiencing that level continuously (at least 3 days)

of global planting, equivalent to 42% of the grain-planted land in Turkey. Despite being the 8th largest exporter of wheat according to USDA figures cited in the TEPGE report, Turkey's self-sufficiency rate on wheat has decreased from 100% to 83.3% as of the 2021/2022 harvesting period. The consumption per capita stands at 179.3 kilograms [33].

The 2022 Ministry of Agriculture report reveals that 78.1% of Turkey's wheat consumption is for food, 12.4% for feed, and 6.6% for seeds. Per capita consumption decreased from 192.8 kilograms in 2019/20 to 176.8 kilograms in 2020/21. Notably, Iraq, Venezuela, Somalia, Benin, and Ghana are the leading export destinations for Turkish wheat, while Russia tops the list of wheat importers [36].

Over the last 60 years, world wheat yield has experienced a remarkable linear increase of about 40 kilograms per hectare per year, suggesting a projection to meet future demand adequately and balance global wheat demand growth [36].

2.1.5 Some Recent Public Incentives on Wheat

For certain agricultural products such as Wheat, Barley, Rye, Oat, Corn, Paddy, Rice, Legumes, Poppy, Hazelnut, Dry Raisin, Dry Fig, Dry Apricot, public subsidies are facilitated through the Turkish Grain Board (TMO), which serves as a market maker. Standards for these subsidies are announced via the TMO website, and applications are collected through the ÇKS system, a central database managed by the Ministry of Agriculture and Forestry, where farmers are required to register. This system enables efficient monitoring of subsidies and evaluation of agricultural policies [47] [25]. Among these grains, wheat holds a significant position, with purchase rates representing approximately 15% of total production on average. This ratio is 5% for barley and 12% for corn [48]. TMO utilizes these mechanisms to balance the market and influence farmers' planting preferences to ensure sustainable and efficient yields in the upcoming years.

In response to the Ukrainian war and to safeguard food security, support was provided to wheat producers, and import barriers were implemented in 2022 and 2023. Import taxes for wheat were raised to 130% during this period [33], potentially impacting wheat yields.

2.2 Barley

Barley ranks as the second most common grain after wheat in terms of cultivated land. Like wheat, it is grown across nearly all regions in Turkey. Barley boasts a higher yield per unit compared to wheat, and the majority of the harvest is consumed domestically within Turkey [4].

To comprehend the relationship between meteorological conditions and barley production, it is imperative to analyze the characteristics of the crop, its growth requirements, and its various applications. Additionally, this section delves into the scale of barley production and the public support mechanisms in place.

2.2.1 Physical Characteristics of Barley

Barley offers several advantages over wheat due to its unique characteristics. Ripening earlier than wheat, barley can effectively shield itself from late-arriving droughts, making it a valuable alternative for land utilization in anticipation of future droughts. Its growth pattern and surface coverage allow it to outcompete wild herbs, reducing soil water loss and promoting efficient water usage. However, while barley may yield higher crop outputs than wheat in dry regions, this is not universally true under all conditions. While it can withstand drought during its generative phase, it may not fare as well during its vegetative phase compared to wheat [23].

Barley, classified as a long-day plant, exhibits adaptability to various day lengths. It is notably prolific in tillering, typically ranging between 5 and 8 tillers per plant. With a plant length averaging between 5 and 15 centimeters, barley grains typically contain 9% to 13% raw protein and 67% carbohydrates. Barley has the capability for self-fertilization, necessitating seed renewal every five years. Among cultivated plants, barley can achieve its highest yields when grown in fertile soil and under suitable conditions. Its relatively shallow root system, reaching depths of 80-90 centimeters, necessitates access to nutrient-rich soil for optimal growth [4].

According to the Agricultural Economic and Policy Development Institute (TEPGE) Product Report published in 2022, barley composition includes 52-72% starch, 9-14% protein, and non-starch polysaccharides such as 4-6% cellulose/lignin, 3-6%

beta-glucan, and 4-7% arabinoksilan [13].

2.2.2 Areas of Usage of Barley

Barley serves primarily as animal feed, with its nutritional value equivalent to 95% of corn's. When intended for animal feed, a higher protein content is preferred. Another significant application of barley is in the production of malt. For this purpose, the barley should ideally feature two lines of white grains and possess a low protein content ranging from 9% to 10.5%. Barley's suitability for malt production is attributed to its husk, which protects the coleoptiles during germination and filtration processes, as well as the firm texture of its grains and its traditional usage. Approximately 90% of malted barley is utilized in the brewing industry, while the remaining portion serves as food substitutes. In Turkey, a considerable portion of harvested barley is allocated to malt production, alongside its usage in other soft drinks [4].

According to the FAO, barley holds significance as a model crop for research in various fields such as plant breeding, genetics, and biotechnology. It is commonly employed as a feed grain, offering a composition primarily consisting of carbohydrates and protein, the ratio of which varies depending on the growing conditions. In developed countries and rural areas, barley straw finds usage as animal bedding and feed. Additionally, barley is often incorporated into flour for bread production due to its superior nutritional value compared to wheat, while being a cost-effective alternative [1].

2.2.3 Required Weather Conditions for Production for Barley

Barley stands out among calm climate crops for its specific soil requirements. Thriving in moderate temperatures, it prefers environments neither too hot nor too cold, with high relative humidity. Ideal conditions for barley cultivation include temperatures ranging between 0 to 20 degrees Celsius and consistent relative humidity levels of around 70-80%. While excessive sunlight and low humidity are detrimental, high humidity benefits barley growth, particularly during the harvest. In arid regions, hot winds like simoom during the blooming stage can significantly decrease yield by adversely affecting fertilization and grain development. Barley roots struggle to access water in low-humidity areas, resulting in premature maturation, weak grains, and reduced yield. Barley's tolerance to low temperatures is limited compared to milling wheat, with many varieties unable to withstand temperatures below -15 degrees Celsius, even without snow cover. Consequently, barley planting areas are comparatively restricted [22].

Barley cultivation in Turkey is classified into three categories based on growth characteristics: "summery," "wintery," and "facultative." These categories correspond to three distinct regions: Mid-Anatolia for wintery, Aegean, Mediterranean, and Southeast Anatolia for summery, and Marmara Region for facultative barley types. Informal breeding studies are conducted according to these regional classifications [23].

Although barley requires less water than wheat, adequate soil moisture is essential for achieving high yields and quality grains. Harvested earlier than wheat, barley has a unique ability to absorb salt from the soil, thus preventing soil degradation and alkalization, thereby preserving land fertility. Poorly ventilated and sandy soils incapable of retaining sufficient water can diminish fertility, as outlined in Ministry of Agriculture guidelines [22].

Barley plays a crucial role in preventing soil degradation, primarily due to its salt tolerance and ability to remove excess salt from the soil. In regions where irrigation is necessary for agriculture, barley serves as a viable alternative crop, according to the Agricultural Economic and Policy Development Institute (TEPGE) Product Report [13].

As highlighted in a FAO study on post-harvest operations, barley growth can be affected by environmental factors. Mild winters may lead to excessive growth and dense canopies, while rainy springs can cause lodging, promoting pest and disease development, ultimately reducing yield. Optimal harvesting occurs when the stem is sufficiently dry to break easily by hand. In regions with higher humidity, moisture meters or a simple hardness test can help determine the ideal harvest time. Harvesting before rainfall is preferred to avoid seed discoloration and yield reduction due to delayed harvests [1].

In Table 2.3, the necessary conditions are summarized based on their purpose of use.

As it can be seen, the conditions are more selective for malting kind of barley. Type of barley which is used for animal feed constitutes 95% of the whole barley production in Turkey and can be produced almost all parts of Turkey since it does not need additional irrigation other than natural one [10].

Table 2.3: Necessary Conditions for Barley Production

Barley Type	**Rainfall**	**Avg Temp/Humidity**
Barley for Malting	Annual: >400 mm	Avg for May: <25 Celcius

2.2.4 Significance of Barley

Barley's adaptability and minimal specific growing requirements make it a preferred crop for cultivation in unfavorable climates and soil conditions worldwide. Its versatility has made it a staple crop for centuries, serving various purposes such as animal feed, food, and a crucial raw material for the malt and beer industry. The FAO study "Barley: Post-Harvest Operations" highlights the remarkable diversity of barley cultivation environments, ranging from 330 meters below sea level near the Dead Sea in the Middle East to altitudes of 4200 meters in the Altiplano and the Andes in Bolivia [1].

The historical roots of barley cultivation trace back to the Middle East, particularly the "Fertile Crescent" encompassing Turkey, Iraq, Lebanon, and Iran. Archaeological evidence suggests that barley was among the earliest cereal crops, with domestication occurring around 17,000 years ago in Egypt's Nile River Valley. Barley holds significant importance in semi-arid regions across Africa, the Middle East, the highlands of Asia, the Andean countries of South America, and certain parts of Asia [1].

Barley's stability against seasonal variations and consistent yield make it a preferred choice for farmers, particularly those in economically challenged regions. Its cultivation provides a safety net against crop failure and low yields, enabling farmers to transition to other crops like wheat with more confidence. Turkey, known for its rich genetic diversity, has been a significant center for barley cultivation, housing various cultural types adapted to different ecological regions. Since the early days of the Turkish Republic, regional institutes have conducted breeding studies to develop genotypes suited for different climatic conditions and agricultural practices. However, certain regions, such as the mid-Anatolian and Southeast Anatolian regions, are not considered suitable for barley cultivation for malt purposes due to inadequate rainfall [10].

The TEPGE report [13] provides USDA figures indicating that approximately 159 million tonnes of barley were planted on approximately 51 million hectares (ha) of land worldwide during the 2020/21 season. Europe accounted for the largest planted area and production of barley. According to data from the Turkish Statistical Institute, approximately 3.2 million ha of land were planted with barley in Turkey in 2021, yielding 5.75 million tonnes of barley production with an average yield of 181 kg per decare. TEPGE also provides an overview of the efficiency situation in Turkey, as shown in Table 2.4.

Table 2.4: Efficiency in Barley Production in Turkey

Year	**Planted Land (thousand da)**	**Production (thousand tones)**	**Efficiency**
2018	26,120	7,000	567
2019	28,690	7,600	571
2020	30,971	8,300	578
2021	31,691	5,750	360
2022	32,175	8,451	436

2.2.5 Some Recent Public Incentives on Barley

The subsidy amounts over the last 5 years are presented in Table 2.5, sourced from the Ministry of Agriculture and Forestry report [14].

Table 2.5: Public Supports for Barley in Turkey

Support Item	**2019**	**2020**	**2021**	**2022**	**2023**
Soil Analysis (TL/da)	0.8	0.8	0.8	1	1
Fertilizer support (TL/da)	8	16	20	46	46
Gas support(TL/da)	19	19	22	75	103
Excessive Payment (TL/kg)	0.1	0.1	0.1	0.1	0.1
Certified Seed (TL/da)	8.5	16	16	50	65

2.3 Corn

Corn, also known as maize (Zea mays), is believed to have originated in the Andean region of Central America, according to studies by the Food and Agriculture

Organization (FAO) [15]. It is a crop that can thrive in tropical, subtropical, and mild climate zones, making it suitable for cultivation in nearly all countries except Antarctica. Corn holds significant socio-economic importance not only in Turkey but also worldwide, owing to its various types and extensive adaptability, leading to widespread cultivation in Turkey [3]. Before delving into the relationship between corn prices and meteorological requirements, it is essential to understand the fundamental aspects of corn cultivation and harvesting.

2.3.1 Physical Characteristics of Corn

Corn, native to the Americas, is primarily composed of starch, protein, and fat, and is rich in vitamins A, E, and selenium [3]. It consists of the seed coat, cotyledon, and germ, each with distinct chemical compositions. During processing, corn is broken down into main starch products and by-products such as corn gluten, corn husks, and corn pulp [21].

Corn can be cultivated between the 58th north and 40th south parallels and up to an altitude of 4000 meters [3]. Nitrogen, phosphorus, and potassium are crucial for corn production. Nitrogen significantly affects both the quantity and quality of corn, with a deficiency leading to poor vegetative and generative development, early flowering and reduced development time. Phosphorus deficiency in the initial growth stage can weaken both the root system and the above-ground portion of the plant, impairing starch, oil, and protein metabolism and decreasing yield quality. Similarly, potassium deficiency can reduce grain quality and make the plant more susceptible to drought [6].

Corn is considered a day-neutral or short-day plant in terms of day length. Spacing and sowing methods depend on factors such as soil fertility and water requirements, aiming to achieve optimum density for light interception and maximum yields. According to the FAO database, the recommended crop plant population ranges from 20,000 to 30,000 per hectare. Row spacing is typically best between 0.6 and 1 meter, with seeds sown at a depth of 5 to 7 centimeters [15].

Corn can grow in most soil types except for those with dense clay or sand. It requires well-aerated and well-drained soil, and waterlogging should be avoided. Ultimately, maintaining soil fertility is crucial for continuous corn cultivation [15].

2.3.2 Areas of Usage of Corn

Corn is a highly versatile and widely utilized crop plant, with over 1000 kinds of processing by-products used in various industries such as food, chemicals and fermentation [21]. Approximately 90% of the world's corn production is utilized for human consumption and animal feed, while the remaining 10% is allocated for industrial purposes [3]. Corn lends itself well to different degrees of processing, including primary and deep processing. Primary processing involves simple procedures such as cleaning, dehydrating, crushing and soaking, while deep processing involves more advanced technological transformations of the final product [21].

According to the Ministry of Agriculture, corn is a crucial raw material for several industries, including starch, syrup, beer, industrialized alcohol and whisky production. Additionally, it plays a significant role in alternative fuel and energy research. Corn silage is extensively used as animal feed, particularly in industrialized animal production systems, to feed livestock such as cows for meat and milk production. Corn silage provides essential protein, minerals and sometimes energy to these animals. Corn oil usage is becoming increasingly common, and corn starch is utilized in paper production to bind layers and enhance food structures. Corn is also employed to reduce dough viscosity, increase gelatinization, and produce specific drug tablets such as aspirin. Furthermore, it finds application in wall woods, insulation and acoustic constructions [6].

2.3.3 Required Weather Conditions for Production for Corn

Corn cultivation requires fertile, well-drained soil with neutral pH levels between 6 and 7. While corn does not have specific soil type preferences, it is sensitive to salinity and high water table levels. FAO categorizes corn as moderately sensitive to salinity, with yield showing an inverse relationship with soil salinity levels [15].

Corn undergoes four critical periods during its growth cycle: seedling emergence (vegetative period), pre-top-tasselling, bottom-tasselling (yield formation period) and filling of corncobs (ripening period). The yield is particularly influenced during the period from two weeks prior to top-tasselling until two weeks after bottom tasseling [3].

According to the Ministry of Agriculture's guidelines for corn farmers, optimal sowing occurs when the soil temperature reaches 12 - 13°C at a depth of 9-10 cm to facilitate germination. Corn requires approximately 500 mm of water during its growing period, with specific water requirements for each month: 75 mm in May, 100 mm in June, 175 mm in July, 100 mm in August, and 50 mm in September [6].

FAO notes that corn can survive and adapt to various climates as long as the mean daily temperature remains above 15°C and frost is not an issue, especially during sowing. Adequate irrigation is crucial for survival when temperatures exceed 45°C. Selecting the appropriate corn variety based on desired outcomes is essential for successful cultivation, ensuring that the growing period aligns with the length of the growing season [15].

The other FAO facts can be summarized as Table 2.6, showing the required temperatures based on the phase of growing of the crop.

Table 2.6: Necessary Conditions for Corn Production

Mean Daily Temperature	**Phase of Growing**	**Days to mature**
>20 Celcius	Early grain	80 to 110 days
>20 Celcius	Medium Varieties	110 to 140 days
>20 Celcius	Grown as Vegetable	15 to 20 days shorter than all above
10 to 15 Celcius	Grown as forage	In case of a problem in seed set and grain maturity under cool conditions the crop results in forage.
Min 10 Celcius, 18 to 20 Celcius optimum	Germination	
<20 Celcius	All	Each 0.5 Celcius results 10 to 20 days extention
1.5 Celcius	All	200 to 300 days

According to the FAO resource, optimal root development in corn occurs when watering is provided at or just after sowing. The required water depletion throughout the growth stages can be summarized as follows: 40% during establishment, 55-65% during vegetative, flowering, and yield formation periods, and up to 80% during the ripening period. In conditions where water availability is limited, such as short rainfall or insufficient irrigation, it is crucial to avoid water deficit during flowering and yield formation stages. Similarly, water supply should be managed carefully during the vegetative period and yield formation period to prevent additional losses [15].

In line with this, the study on corn farming conducted by the Directorate of Trakya Agricultural Research Institute provides specific recommendations for sowing dates in different regions of Turkey. Sowing can commence in Thrace after April 25th, in southern regions at the beginning of April, and in Central Anatolia starting from April 25th. However, the best practice is to begin sowing after the last frost dates in each region to mitigate the risk of frost damage to young seedlings. In Turkey, corn sowing should conclude by the first weekend or mid-second week of May to avoid flowering during hot and dry periods, which could result in reduced fertility. Late sowing may also lead to ripening and harvesting coinciding with rainy periods, impacting crop quality [3].

2.3.4 Significance of Corn

Corn is one of the few crops that are raised for thousands of years. As stated in the study of Directorate of Trakya Agricultural Research Institute, the main land is the continent of America, is known to be disseminated from there. According to the archeological studies performed in New Mexico, USA, in the shelters and caves some crop and corncobs were found to be 5000 years-old. Samely, Mexicocity findings shows 7000 yo cornflower dusts under 50-60 m deep. All archaeological findings result in the fact that the corn plant has a past of 8000 and 10000 years [3].

The continent of America had already started to raise corn in several areas by the time it was explored. The types were horse teeth corn, hard corn, floury corn, sweet corn and popcorn. Even Aztecs were worshipping many Corn Gods and pray for a better harvest since those were the main nutrient of the local tribes around Mexico, mid and south of Americas. Similarly in the mythology of Native Americans the corn took its place. The Spanish and British explorers learned how to raise corn and where to use them from the Native Americans. [3].

Christophe Colomb has introduced corn to Europe by bringing them in his pocket to Spain. Following that, Portugal, France and Italy started to raise it then Southeast Europe and North Africa farms have widely been sown corn. Portuguese sailors then brought corn to west coasts of Africa in the beginning of the 16th century, then India and China soon spread to the whole Asia. With its rapid breeding and high yield potential corn plant has easily surrounded every region it entered and even replaced the main crop in that land. Corn has entered Turkey from North Africa, towards Egypt and Syria, which can be understood from the Turkish word “Mısır” same as the Turkish name of the country Egypt [3].

In Turkey, corn, 83% of which is used by feed industry, is mostly raised in the Konya (15%), Şanlıurfa (19.2%) and Adana (10.1%) covering the whole 6 million tons of Turkey’s production with a 44.3% ratio, as per 2020 data given by Agricultural Products Market report. Same rank is maintained for 2021 and 2022, the ratios respectively are Konya (18.7%, 24%) Urfa (12.1%, 10.2%) and Adana (12%,10.5%). The cultivated land ratio increased up 9.1 million tons. Konya raised its production level above 2 million tons constituting 24% of the whole country in 2022. Ankara has a production of 63 thousand tons of corn as of 2022 [41] [42] [43].

Global corn production on the other hand has exceeded 1.1 billion tons in 2020/21 and in 2021/22 1.16 billion tons. It is most produced in USA however the largest field of corn belongs to China. The third place in corn production is Brasil with 11.5 billion tons. Brazil is the largest exporter, while China is the largest importer according to 2022/2023 harvest year figures [41] [42] [43].

2.3.5 Some Recent Public Incentives on Corn

Ministry of Agriculture and Forestry implements various subsidies, with one of the most significant impacts for corn being the "excessive payment" subsidy. Given that corn cultivation typically requires extensive irrigation, there’s a focus on ensuring efficient water usage. In provinces like Konya, Karaman, and Hatay, where water resources are scarce, only farmers utilizing drip irrigation systems are eligible for public subsidies. This policy aims to encourage water-efficient farming practices and mitigate the strain on limited water sources in these regions [42].

2.4 Phenological Effects

Phenology refers to the natural life cycle of living organisms in general, providing insight into the seasonality of specific plants. According to the Phenological Atlas study prepared by the Meteorological Service of the Turkish State, it is the science that measures the timing of life cycles for plants, animals, and microorganisms, and determines the effects of environmental conditions on these timings. The schedule of phenological phases can vary between years due to meteorological models and climate changes. In essence, seasons are more dependent on phenological cycles than on the calendar itself [38].

Phenology encompasses not only the effects of temperature but also factors such as altitude, humidity, solar absorption, and rainfall. Therefore, it serves as an interdisciplinary field bridging climatic and agricultural sciences, closely linked with geography and biology [38].

Phenological observations play a crucial role in selecting the most suitable crop varieties to adapt to climate conditions, as well as in their maintenance and breeding. Similarly, for preventing frost damage, information on the sensitive phases of specific crops is essential, obtained by comparing phenological observation dates. This enables rational agricultural practices, such as determining the optimal timing for sowing, harvesting, and applying pest control measures, thereby maximizing yield efficiency. [38].

Table 2.7: The Phases of Phenological Observations of Grains

No	**Phase**	**Wheat and Barley**	**Corn**
0	Sowing	Date of sowing the seeds in the soil	Same
1	Germination	The first leaf to go above the ground	Same
2	Leaf line	The third leaf to reach 1 cm	The ninth leaf to reach 1 cm.
3	Tillering	The first body knot to be above the ground by 1-2 cm	Same
4	Heading	The phase that plant grows rapidly	The corncobs emerge from the leaf collars
5	Earing	The cover of the middle leaf grows into ear and it exposes	Same, named as tillering
6	Flowering	The phase that flowers blossom and the pollens spread	Same
7	Ripening	The feekes turn into its typical color.	The grains in the middle of corncobs turn into its typical color The grain is as soft as dough. Some part of the plant (the below leaves and the tassels) are dry and yellow as can be observed by shaving the covers of the corncob.
8	Harvesting	The grain gets solid and the feekes to be mature enough to be harvested	The most of the leaves are yellow and dry.

According to the studies held in 2014 by Research Department of Agricultural Office in Meteorological Service of Turkish State on the phenological effects [38], the phases of phenological observations of Wheat, Barley and Corn are given in Table 2.7.

Presented next are isophane maps, also known as phenological atlases, for each selected grain. These maps utilize color schemes to represent specific time periods, such as sowing, earing, and harvesting, as indicated by the legend. Understanding these time intervals is crucial for the financial model in this study, as they form the basis of each grain's life cycle. Additionally, the maps include information on precipitation and dew point degrees, allowing for correlations with annual yield [38].

The color scheme in Figure 2.1 shows the life cycle of **Wheat** in Turkey.

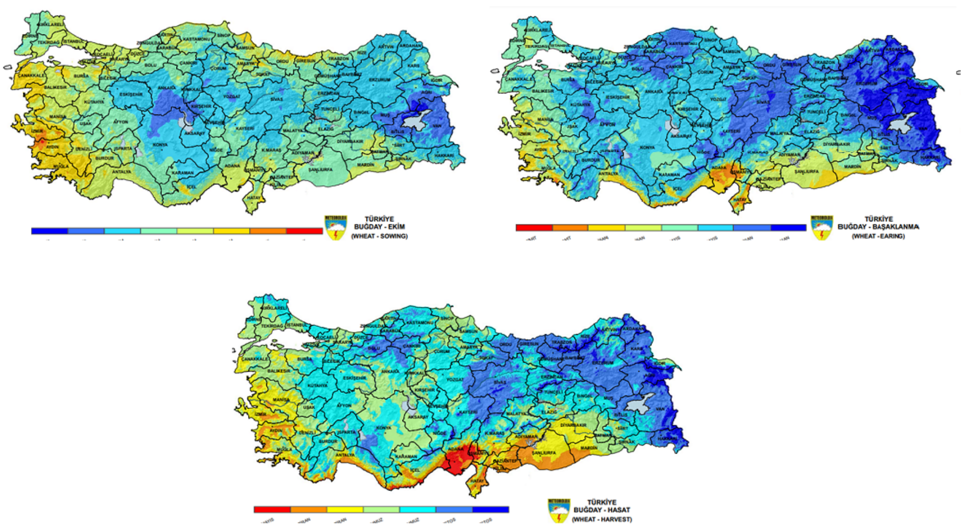

Figure 2.1: Wheat Isophane Maps in Turkey

- The map on the left-hand side displays the timing of wheat **sowing** across Turkey. The color palette represents the following date ranges: Sep 1-15; Sep 16-30; Oct 1- 15; Oct 16-31; Nov 1-15; Nov 16-30; Dec 1-15; Dec 16-31. Observing the map, wheat sowing initiates earliest in the eastern region and concludes lastly in the western region of Turkey.
- The map on the right-hand side displays the timing of wheat **earing** across Turkey. The color palette indicates the following date ranges: Mar 1-15; Mar 16-31; Apr 1- 15; Apr 16-30; May 1-15; May 16-31; Jun 1-15; Jun 16-30. It can be observed that wheat begins earing by the end of March in the southern part

of Turkey and continues until the second half of June in highlands, particularly in the northeastern region.

- The map at the bottom displays the timing of wheat **harvesting** across Turkey. The color palette represents the following date ranges: May 16-31; Jun 1- 15; Jun 16-30; Jul 1-15; Jul 16-31; Aug 1-15; Aug 16-31. Notably, wheat harvesting commences earliest in the southern region in the second half of May and concludes lastly in the eastern region of Turkey by the second half of August.

The color scheme in Figure 2.2 shows the life cycle of **Barley** in Turkey.

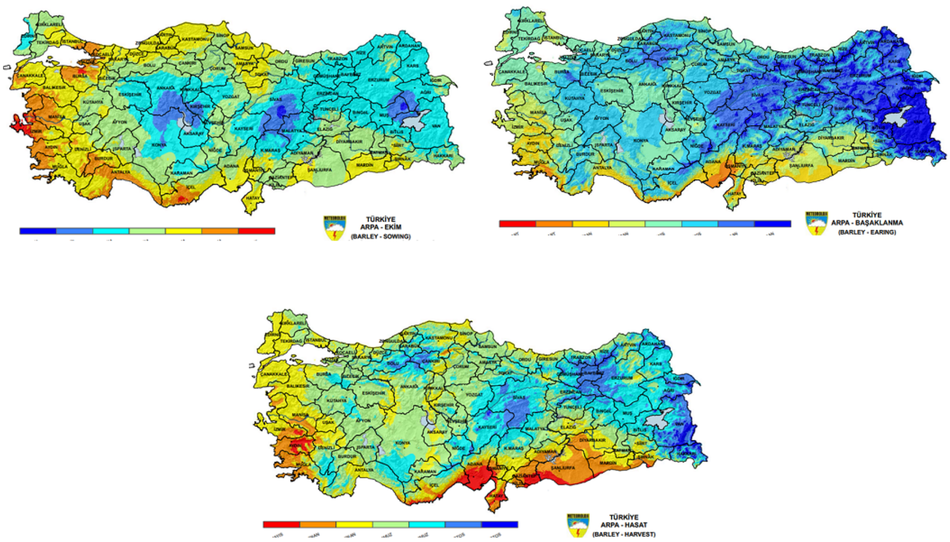

Figure 2.2: Barley Isophane Maps in Turkey

- The map on the left-hand side displays the timing of barley **sowing** across Turkey. The color palette indicates the following date ranges: Sep 1-15; Sep 16-30; Oct 1- 15; Oct 16-31; Nov 1-15; Nov 16-30; Dec 1-15. Observing the map, barley is first sown in the eastern part, including Sivas and Muş, around the first half of September and lastly in the western part of Turkey, including İzmir and Bursa, as well as in the southern part, around İçel, around the first half of December.

- The map on the right-hand side displays the timing of barley **earing** across Turkey. The color palette represents the following date ranges: Mar 1-15; Mar

16-31; Apr 1- 15; Apr 16-30; May 1-15; May 16-31; Jun 1-15; Jun 16-30. It can be observed that barley begins earing first in the southern part, in Adana, in March and concludes lastly in the eastern region of Turkey, in the second half of June.

- The map at the bottom showcases the timing of barley **harvesting** across Turkey. The color palette indicates the following date ranges: May 16-31; Jun 1- 15; Jun 16-30; Jul 1-15; Jul 16-31; Aug 1-15; Aug 16-31. Notably, barley is first harvested in the southern region, as well as around İzmir and Aydın in the west, in the second half of May and lastly harvested mostly in the eastern region in the second half of August.

The color scheme in Figure 2.3 shows the life cycle of **Corn** in Turkey.

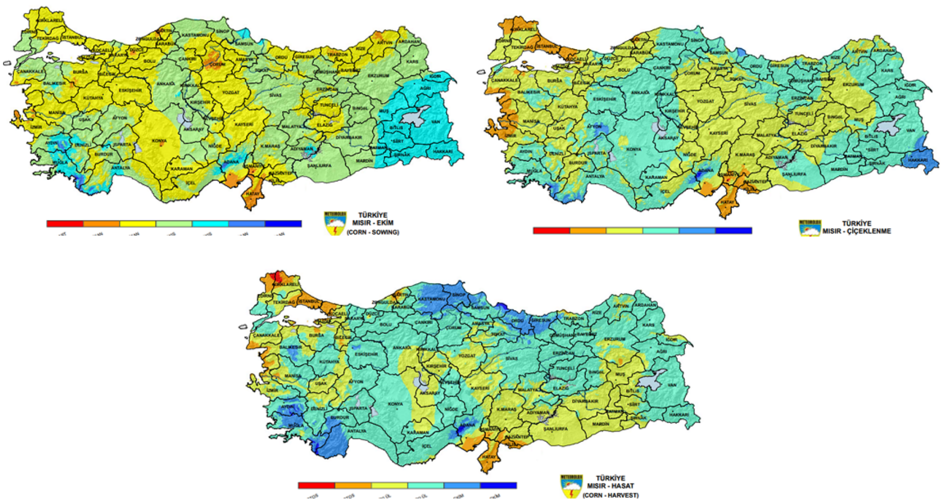

Figure 2.3: Corn Isophane Maps in Turkey

- The map on the left-hand side illustrates the timing of corn **sowing** across Turkey. The color palette represents the following date ranges: Mar 16-31; Apr 1- 15; Apr 16-30; May 1-15; May 16-31; Jun 1-15; Jun 16-30. It can be observed that corn is sown starting from the first half of April in several parts of the Black Sea region, some points in the Aegean, and the southern part of Turkey until the first half of June, especially in the eastern region.

- The map on the right-hand side displays the timing of corn **flowering** across Turkey. The color palette indicates the following date ranges: Jun 16-30; Jul 1-

15; Jul 16-31; Aug 1-15; Aug 16-31; Sep 1-15. It can be seen that corn begins flowering first in Osmaiye in the second half of June and concludes lastly in the southeast, including Hakkari and the highlands of Adana, in the first half of September.

- The map at the bottom displays the timing of corn **harvesting** across Turkey. The color palette corresponds to the following date ranges: Aug 1-15; Aug 16-31; Sep 1-15; Sep 16-30; Oct 1- 15; Oct 16-31. Observing the map, it is evident that corn harvesting commences earliest in the northwest region, specifically in Kırklareli and Edirne, during the first half of August. Conversely, the last areas to harvest corn include the Black Sea region, Samsun, and the highlands of Adana in the southern part of Turkey, occurring in the second half of October.

2.5 Insurance Policies

Lastly, essential information of crop insurance applications in Turkey is explained in this section to have an understanding on its significance in the country's agricultural sector. Even though the model does not directly incorporate insurance-related inputs, its outcomes could be valuable for insurance companies in modeling their insurance risks.

When deciding which crop to produce, factors such as weather conditions, land characteristics, political incentives, and insurance coverage percentage are crucial considerations. Given that the farmer's entire year of effort culminates in a single harvest, insurance becomes a necessity rather than an option. The premiums paid for insurance are considered as production inputs and can influence prices. Unusual weather conditions lead to higher premiums due to increased risk of no return [40].

Agricultural insurance was first legislated by Law No. 5363, the Agricultural Insurance Law, published in the official gazette on June 21, 2005. Public authorities support the Crop Insurance system by covering 50% of the premiums stated in the policy. TARSIM (Agricultural Insurance Pool) is the main institution responsible for insuring risks, covering only farmers registered and providing updated information in the Farmers Registration System (FRS). According to TARSIM's guidelines, crop quantity losses due to hail, storm whirlwinds, fire, earthquakes, landslides, and floods are covered, with the option to cover certified seeds and losses caused by hot weather [40]

The insurance system operates as follows:

- When risk is realized; the policyholder/insured notifies TARSIM within 10 days for frost risk and 15 days for other risks, directly or through agencies.
- Loss adjustments are made by the loss adjusters appointed by the TARSIM.
- Following the completion of the loss file, finalized indemnity amounts are paid within 30 days at the latest.
- Indemnities are paid at the harvest date. However, in case of a total loss, within the principles of insurance, indemnity can be paid before the policy termination date.
- When the loss happens in an early growing phase so that re- sowing /re-planting of the same crops possible, sowing/planting and care expenses incurred by the insured up to the date of loss can be paid.

According to 2020 and 2021 figures, published in 1st July 2022 bulletin of TARSIM (no. 45521), the largest indemnity payment in 2021 was due to frost, amounting to 988,907,704 TL, while the largest payment in 2020 was due to hail damage, totaling 479 million TL. Public inducements on insurance premiums were increased by 49.1 % between 2020 and 2021. On the other hand the number of policies were increased from 1598 to 2518 and the total amount was increased from 30.3 billion TL to 124.4 (quadrupled) between years 2017 and 2021. Furthermore, the farmland area (including greenhouses) increased by 7.1%, resulting in 28,798,222 da [52]. On the other hand, thinking about the cultivated land of wheat only which is 6,6 mio ha in Turkey [33], crop insurance sector may be concluded to have still a long gap of coverage to fill.

In conclusion, the model's outputs, although not directly incorporating insurance-related inputs, can be instrumental for the crop insurance sector in Turkey. By providing accurate forecasts and data-driven insights, insurance companies can better assess and model risks, leading to more effective and responsive insurance products.

This, in turn, benefits farmers by offering more tailored insurance coverage, thereby enhancing their financial security and resilience against climate-induced uncertainties. Thus, leveraging such advanced models in the crop insurance sector is not only beneficial but also a necessary step toward a more robust agricultural framework in Turkey.

2.6 Conclusion

This chapter provides an explanation of the selection criteria for the chosen crops, namely wheat, barley and corn, to justify their inclusion in the model studied in this thesis study. By examining the characteristics of these crops, including their grains, growing conditions and timelines in specific regions, a comprehensive understanding of these grains has been developed. Considering their historical background both globally and in Turkey, it becomes evident that wheat, barley and corn are majorly produced and consumed grains, essential for sustainable production and trading.

Moreover, the chapter delves into phenological information, which offers insights into the timing of weather effects, as seasons are often dictated by phenological cycles rather than the calendar itself. Utilizing historical data on climate variables such as temperatures (with a focus on dew point for humidity levels) and precipitation relative to yield, the first model is constructed based on phenological cycles (see Section 3.1.1.3).

Additionally, the section addresses the significance of crop insurance coverage, providing background information to underline the potential use of output data in the risk modeling of crop insurance companies.

Furthermore, the chapter elucidates the motivations behind public incentives and insurance policies for farmers, highlighting the physical conditions necessary for crop growth, including the nature and reproductive potential of seeds, phenological effects across the country's landscape and the timing of weather factors such as rainfall and temperature which are the most essential factors in the lifecycle of the grains as pointed previously in this chapter, also in Section 3.1.1.1 that they are not only pivotal factors but also measurable so that instantly recordable. These aspects collectively inform the selection of the dataset for the model (see Section 3.1).

In the subsequent chapter, a financial model is developed to analyze the impact of climate factors on future crop prices, structured in two layers. The first layer is based on yield as the dependent variable, with total precipitation and average dew point based on the essential periods of the grain lifecycle, of the selected province serving as independent variables. The second layer utilizes annual yield as an independent variable, along with net import yield and global prices, to link the dependent variable of local price. The overarching objective is to achieve an accurate estimation of future prices for wheat, barley and corn by leveraging weather forecasts on total precipitation and dew point. Ultimately, the aim is to establish a base for generating a future price index based on future weather forecasts, as outlined in the following chapter.

CHAPTER 3

THE FINANCIAL MODEL

Based on the comprehensive understanding of the characteristics and cultivation requirements of wheat, barley, and corn, as well as their historical significance globally and within Turkey, including considerations of public incentives and insurance protections, a clearer perspective on these grains has been established. Leveraging this knowledge, the financial model is designed to employ a time series analysis using historical data encompassing climate variables such as temperatures (specifically dew point for humidity levels) and total precipitation, correlated with yield. This correlation is then further examined in conjunction with domestic prices, accounting for inflation effects, as well as global prices and import/export figures.

Consequently, two consecutive models are formulated: the first model focuses on yield as the dependent variable, with total precipitation and average dew point of the selected province serving as independent variables. Subsequently, the second model utilizes annual yield as the independent variable, alongside net import yield and global prices, to predict the dependent variable, local price. This approach aims to forecast the impact of specific climate factors on local prices.

The overarching objective is to achieve precise estimations of future grain prices by leveraging weather forecasts for total precipitation and dew point. Ultimately, this endeavor seeks to generate a future price index based on anticipated weather conditions, facilitating informed decision-making in commodity markets.

In this chapter, the groundwork for the analysis is laid out, beginning with the preparation of the dataset and its subsequent grouping. The rationale behind the chosen methodology is elucidated, providing insight into the decision-making process driving the model's construction.The ensemble machine learning method is detailed, highlighting its suitability for the analysis and its application in the context of the study's objectives. The chapter further delves into the procedural steps involved in implementing the model across two layers, utilizing tree-fit and also simple regression approach to offer a contrasting perspective on the modeling approach.

Following the execution of the models, the results are systematically compared and interpreted. Descriptive statistics are employed to provide a comprehensive understanding of the findings, shedding light on the efficacy and performance of each model. Through this comparative analysis, insights are gleaned into the predictive capabilities of each approach and their respective strengths and limitations.

3.1 Data Set

A panel dataset is constructed, as depicted in Table 3.1, with the belief that assembling multiple layers of data will yield a more accurate model. The model is structured in two phases: Phase 1 examines the impact of weather conditions on yield, while Phase 2 aims to predict the effect of yield (alongside global price and net import) on price. This sequential approach enables the prediction of the Phase 2 equation (local price) using the results of the Phase 1 equation (local yield). Consequently, the ultimate objective of forecasting prices based on climate factors such as total precipitation and dew point can be realized.

3.1.1 Data Selection Criteria

3.1.1.1 Weather Input Selection

Within the model examining the relationship between yield and meteorological events, the analysis focuses on two pivotal factors among the measurable and monthly recorded units: natural irrigation and temperature. Instead of solely considering rainy days or snowfall, the dataset collects total precipitation data, reflecting the cumulative natural irrigation essential for the grains from planting beneath the soil to maintaining necessary moisture levels during germination and subsequent stages until harvest. Addi-

Table 3.1: Financial Model Components

	Phase 1	Phase 2
Aim of the Model	Yield Prediction based on Total Precipitation and Dew Point	Price Prediction based on Yield, Net Import and Global Price
Dependent Variable	Yield	Local Price
Independent Variables	Scores made from: • Monthly Total Precipitation • Monthly Dew Point	• Periodical Yield • Periodical Net Import • Global Price at Period end
Equations set	• Per Grain – Wheat – Barley – Corn • Per Province – Konya – Polatli – Yozgat – Adana – Sanliurfa	• Per Grain – Wheat – Barley – Corn • Per Period – Period 1: January-February-March – Period 2: April-May – Period 3: June-July – Period 4: August-September – Period 5: October-November-December
Years	2000-2022	2010-2022

tionally, temperature figures are converted into dew point values (2-meter dewpoint temperature), which represent the humidity saturation level, indicative of the moisture content in the air. Essentially, the dew point signifies the temperature at which the air must be cooled to achieve saturation, with higher values indicating greater atmospheric moisture.

3.1.1.2 Grain Type and Province Selection

The selection of grains is based on their significant production volumes and trading ratios, which align with the three criteria used for determining the inclusion of products in TMEX commodity indices [46]:

- **Production Quantity:** The average production and import quantity over the last 3 years must exceed 1 million tons.

- **Trading Volume:** The average annual trading volume over the last 3 years must exceed 100 thousand tons, with an average daily trading volume exceeding 500 tons.
- **Days of Trading:** Over the past year, the number of trading days must be at least 60% of the total business days in a year.

In TMEX, all three grains have their own indices which are Wheat Index, Barley Index and Corn Index [46], justifying the significance of these grain types to be included in the model in this thesis study.

The selection of provinces aimed to encompass the primary production areas in Turkey, with wheat being the focal point due to its significance as elaborated in 2, since it is the main ingredient for bread and also has a vast area of usage. Konya, Urfa, Ankara, Diyarbakır, Tekirdağ, Adana, and Yozgat emerged as the major wheat-producing regions, accounting for 8.86%, 4.96%, 4.94%, 4.61%, 3.73%, 3.67%, and 3.46% of the total average production between 2000 and 2022, respectively, based on TURKSTAT data. Among these high-yield regions, two groups were formed based on their closely aligned climate and geographical characteristics: the first group comprises Konya, Polatlı (a significant wheat-producing province in Ankara), and Yozgat, while the second includes Adana and Urfa, focusing on durum wheat production. These provinces were also chosen for the study of other grains such as Corn and Barley, as they remain pivotal production regions in Turkey, allowing for meaningful comparisons. Together, these five provinces (with Ankara's yield representing Polatlı) collectively contribute 26% of the nation's wheat production, 28% of barley production, and 35% of corn production on average between 2004 and 2022, according to TURKSTAT statistics [49].

Once the provinces are determined, next the coordinates of each province are taken as follows (latitude and longitude points are provided in brackets respectively), retrieved with a grid of [0.1, 0.1].

- Konya: [39., 31.5, 37., 34.] [1] as the matrix size would be 26x21 making 546 dots per month in 23 years.

[1] The starting and ending points for latitude and longitude, respectively, corresponding to other provinces accordingly.

- Polatlı: [40., 32, 39., 33.] as the matrix size would be 11X11 making 121 dots per month in 23 years

- Yozgat: [41., 34., 39., 36.] as the matrix size would be 21x21 making 441 dots per month in 23 years.

- Adana: [39., 35., 36., 37.] as the matrix size would be 21x31 making 651 dots per month in 23 years.

- Urfa [38., 38., 36., 40.] as the matrix size would be 21x21 making 441 dots per month in 23 years.

3.1.1.3 Phase 1_Period Selection

According to the isophane maps depicting grain cultivation in Turkey, as illustrated in the figures under section 2.4, temperature and precipitation play crucial roles during specific periods, as outlined in Table 3.2, across the provinces where they are cultivated.

Table 3.2: Essential Periods on the Life of Grains

Grains	**Provinces**	**Sowing & Germination**	**Earing**	**Harvesting**
Wheat	Konya, Yozgat, Polatlı	Sep 16-30; Oct 1-15	May 16-31; Jun 1-15	Jul 1-15; Jul 16-31
	Adana, Urfa	Sep 16-30; Oct 1-15	May 16-31; Jun 1-15	Jul 1-15; Jul 16-31
Barley	Konya, Yozgat, Polatlı	Sep 16-30; Oct 1-15	May 16-31; Jun 1-15	Jul 1-15; Jul 16-31
	Adana, Urfa	Sep 16-30; Oct 1-15	May 16-31; Jun 1-15	Jul 1-15; Jul 16-31
Corn	Konya, Yozgat, Polatlı	Apr 16-30; May 1-15	Jul 16-31; Aug 1-15	Sep 1-30
	Adana, Urfa	Apr 16-30; May 1-15	Jul 16-31; Aug 1-15	Sep 1-30; Oct 1-15

Additionally, the precipitation before earing has always been crucial for influencing the quality of the yield. Increased snowfall provides a natural cover for the grain, keeping it warm under the soil and facilitating gradual natural irrigation. Therefore, in the model, precise attention is given to the sowing period, while an earlier timeframe is considered for the remaining stages of the grain's life. Consequently, the data collected from ECMWF[12] is grouped into three periods, reflecting the lifespan of a grain from the day it is sown until harvest. These crucial periods, where precipitation and dew point play a significant role, include: (1) sowing, when the seeds meet the soil, (2) the period just before or during the early phase of germination, and (3) the time just before harvest. For the regions of Konya-Polatlı-Yozgat and Adana-Urfa, the specific periods are delineated as shown in Table 3.3, determined based on isophane maps outlined in the phenological atlas of the Turkish State Meteorological Service under Section 2.4. These periods span two to three months, varying according to the vastness and diverse characteristics of each region.

Table 3.3: The Periods Based in Phase 1

Grains	**Provinces**	**Period No 1**	**Period No 2**	**Period No 3**
Wheat	Konya, Polatli, Yozgat	Sep (9) - Oct (10)	Feb (2)- Mar (3) – Apr (4)	May (5) – Jun (6)
	Adana, Urfa	Oct (10) – Nov (11)	Jan (1) - Feb (2)	Mar (3) – Apr (4)
Barley	Konya, Polatli, Yozgat	Sep (9) - Oct (10)	Feb (2)- Mar (3) – Apr (4)	May (5) – Jun (6)
	Adana, Urfa	Oct (10) – Nov (11)	Jan (1) - Feb (2)	Mar (3) – Apr (4)
Corn	Konya, Polatli, Yozgat	Apr (4) - May (5)	Jun (6)	Jul (7) - Aug (8)
	Adana, Urfa	Apr (4) - May (5)	Jun (6)	Jul (7) - Aug (8)

3.1.1.4 Phase 2_Period Selection

The periods in Phase 2 are determined by considering the month when the grain is harvested, which is also regarded as the start of the storing period [45]. Additionally, the future quotation value dates provided in basic future commodity markets such as NMEX [29]and CBOT [9] [8] were taken into account, as shown in Table 3.4.

Table 3.4: Storage Periods & Some Examples of Future Contract Value Dates

Contract	**Value dates (Expiration of the contract)**	**Commodity Exchange**	**Storage Period Start Date**
Wheat	Mar, May, Jul, Sep, Dec	CBOT (a.k.a. CME)	May 15th
Barley	Apr, May, Jun, Jul	NMEX (a.k.a.NCDEX)	June 1st
Corn	Mar, May, Jul, Sep, Dec	CBOT (a.k.a. CME)	August 1st

To maintain consistency with Wheat and Corn, the periods for Barley are also delineated as January-February-March (Period 1), April-May (Period 2), June-July (Period 3), August-September (Period 4), and October-November-December (Period 5).

Consequently, all variables are categorized according to these periods.

3.1.2 Data Time Intervals and Sources

The complete details of the dataset along with their respective sources are outlined in Table A.1 in Appendix. Prices are collected on a daily basis but are aggregated monthly for consistency. Dew point and total precipitation data are recorded monthly, while net imports and production figures are reported annually, given that harvesting occurs once per year.

3.2 Methodology

The dependent and independent variables, as outlined in the preceding section and detailed in Table 3.1 are utilized to establish two equation forms across two distinct phases.

Phase 1 establishes the relationship between weather conditions -specifically, 2-meter precipitation and near-surface air temperature (represented by dew point figures)- and yield for selected grains in each of the chosen provinces. The selected provinces include Konya, Polatlı, Yozgat, Adana, and Şanlıurfa, while the selected grains are Wheat, Barley, and Corn. This results in 15 combinations, each corresponding to a unique province-grain pair. Consequently, 15 equations are formulated for the period spanning 23 years, from 2000 to 2022, as depicted in Equation 3.1.

$$Y_{G,Province,year} = \beta_0 + \beta_1 * X_{G,P,1} + \ldots + \beta_{(n-1)} * X_{G,P,(n-1)} + \beta_n * X_{G,P,n} + \epsilon \quad (3.1)$$

In the context of the model, G represents the grain type, while P denotes the province and ϵ, the error term. The variable $X_{G,P}$ signifies the corresponding score derived from total precipitation and dew point at the province's coordinates. The coefficient β accompanies the number of scores n, which are determined individually throughout the model's course, as detailed in Section 3.2.2.

In Phase 2, the aim is to establish the relationship between local yield, net import, global price and local price for selected grains within predetermined periods. These periods are based on the delivery dates of future contracts in commonly traded markets (see Section 3.1.1.4). With 3 grains and 5 periods, this results in 15 equations per year, which are then applied over a span of 13 years between 2010 and 2022, as illustrated in Equation 3.2.

$$LP_{G,Per,year} = \beta_0 + \beta_1 * MX_{G,Per,year} + \beta_2 * LocalYield_{G,Per,year} + \beta_3 * GP_{G,Per,year} + \epsilon \tag{3.2}$$

In the equations, G still represents the grain type, while LP now denotes the local price, MX net import, GP global price and ϵ, the error term. The coefficient β represents the coefficients accompanying the model inputs, with calculation details elaborated in Section 3.2.2.2.

The main idea is to be able to predict LP , by inserting the output Y of all provinces obtained in Equation 3.1, derived using total precipitation and dew point, into Equation 3.2 as $LocalYield$. This interlinking of the two equations demonstrates how climate indicators can be utilized for local price prediction.

For both phases, two distinct approaches are employed. On one side, the "ensemble tree" method in machine learning (as referred by "tree-fit" in this paper) is utilized, with the main principles elaborated in Section 3.2.1. On the other side, classical Simple Regression Prediction is conducted to compare and interpret the efficacy of the results.

3.2.1 Regression Analysis Methods

To establish a meaningful relationship between the dependent and independent data, regression models were examined. Given the limited expansion of the data, making substantial predictions was challenging, prompting the exploration of different new algorithms. The observed data exhibited seasonality, as seen in total precipitation, dew point, and yield. However, price data did not follow a seasonal or linear pattern, as it is influenced by a wide range of factors, from global crises to consumer preferences.

Total precipitation and dew point data would not be assessed as the single factor for yield and more precise price predictions would require additional components such as freight costs, national prices, and yields in other provinces. Since the representation power of the inputs are not high as desired, an ensemble method among continuous models was chosen. This approach leverages decision trees to produce a stronger *ensemble* output [11]. Given the diverse characteristics of the data, a model that could combine all inputs and interpret them homogeneously, filling in the blank parts with its own intelligence, was deemed suitable. These considerations led to the search for an appropriate machine learning system to work with.

Machine learning systems use algorithms that can *learn* from data directly, rather than relying on predetermined equations. These systems make predictions through computational methods applied to a given set of data. Typically, these models are built using supervised machine learning algorithms. The algorithm assesses the relationship between actual inputs and outputs, training the machine learning model to make reasonable predictions for new data sets [26].

As stated in "Ensemble Methods in Machine Learning" by Dietterich, ensemble methods build classes and make classifications on new data points based on their predictions. These methods, such as the Bagging algorithm, even include correcting errors and tuning outputs accordingly [11]. Dietterich's study demonstrates that ensemble methods are more effective than single methods, with ensemble decisions being more accurate than individual classifiers. This statement is supported by a 2022 study on machine learning and ensemble learning, which found that ensemble methods generalize better and have a lower risk of overfitting [31].

In ensemble learning, predictions may be based on either classes or individuals. Ensemble decisions are composed of individual decisions, and by combining them, new decisions can be made while minimizing errors. In other words, the model's estimations are based on the majority estimation, combining individual estimates [5].

3.2.1.1 Applied Methods in the Study

When examining the framework for ensemble learning, various types can be identified. The first type focuses on data classification, while the second type primarily involves data regression. The models used for classification are explained below for Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN):

- Naive Bayes is known for its simplicity, understandability, and ease of application, but it is typically used in text classification. It calculates the probability of a sample belonging to a target attribute's class, making it irrelevant for our numerical dataset [31] [26].

- Support Vector Machine (SVM) addresses both binary and multiclass classification problems by finding the closest examples of the classes while maximizing the perpendicular distances to the separating surface. It essentially classifies data by identifying the linear decision boundary (hyperplane) that distinguishes data points of one class from those of another [31] [26].

- K-Nearest Neighbor (KNN) classifies objects based on their proximity to other objects. While KNN is easy to develop, it requires significant memory space, and its processing load and cost increase with the dataset's size. Its performance is influenced by parameters such as the number of k neighbors. KNN predicts by categorizing objects based on the classes of their nearest neighbors, assuming that nearby objects are similar [31] [26].

These models would not be suitable for the dataset used in this study, since the weather inputs exhibit seasonal and linear trends, making clustering unnecessary. The classifiation methods of machine learning are not found applicable for our main dataset which includes only two blocks (precipitation and dew point) and classifying into other segments is not preferred for the ease of work [31] [26].

Some examples of machine learning models used for regression include [31] [26]:

- Logistic Regression (LR): This statistical method predicts binary classes. Logistic Regression estimates the probability of an outcome that can only have two values, based on one or more numerical and categorical predictors. Linear Regression is unsuitable for binary systems like yes/no, as it can predict values outside the 0 and 1 range. Logistic Regression, on the other hand, produces a logistic curve limited to values between 0 and 1.

- Decision Trees: A decision tree predicts responses to data by following the decisions from the root to a leaf node. It consists of branching conditions where a predictor's value is compared to a trained weight. The training process determines the number of branches and the values of weights. Additional modifications, or pruning, can simplify the model.

- Ensemble Trees: Ensemble methods combine several weaker decision trees into a stronger ensemble. A bagged decision tree consists of trees trained independently on bootstrap samples of the input data. Boosting involves iteratively adding and adjusting the weight of weak learners, emphasizing misclassified observations or fitting new learners to minimize the mean-squared error between the observed response and the aggregated prediction of all previously grown learners.

Considering the varied nature of the data with respect to seasonality and linearity, rather than classification techniques, various regression approaches are found suitable. Although there is the potential to incorporate numerous inputs for price prediction, this research primarily focuses on precipitation and dew point. Consequently, **Ensemble Trees** method (referred by **"Tree-Fit"** term in this study) is regarded as more appropriate compared to other techniques. The command **"fitrensemble"** is used in MATLAB which is the primary software in the study. This command optimizes the result by estimating the generalization error of an ensemble of boosted regression trees [26].

Regression analysis is used to model the linear relationship between two or more variables. When the dependent variable is continuous, a linear regression model is often used, assuming the error terms have a normal distribution [24]. Ensemble trees method classify the input (X) - output (Y) combinations of a function that is not yet known.

In this study, inputs are vectors with a set of data inside (period-wise or grain-wise), and Y is in the form of a line rather than a classifier. Since the Y vector is predicted using predictor data in the matrix X, the appropriate command in MATLAB "`fitrensemble`" is used. This command optimizes the result by estimating the generalization error of an ensemble of boosted regression trees. This method is among the machine-learning regression methods, specifically the **ensemble trees** model in supervised learning (using both input and output data to form predictive models), as

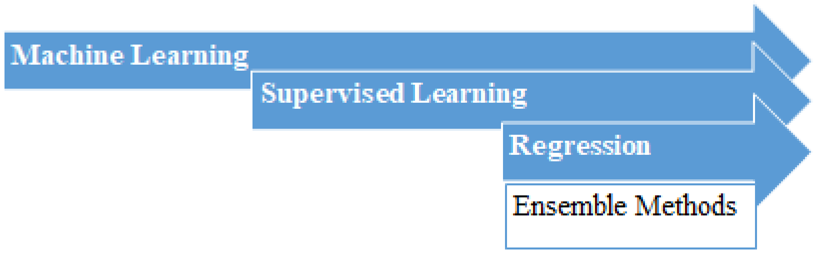

Figure 3.1: Ensemble Method Applied in the Study under Machine Learning

shown in Figure 3.1 [26].

In the ensemble tree method used in the model, the gaps in the data are filled by artificial intelligence to form a strong decision tree, as suggested by the term "ensemble." The decision trees are essentially composed of individual trees trained independently using input samples. According to the "MATLAB, how it works?" page [26], the weight of data lacking desired representation power is incrementally adjusted in such a way that the mean-squared error between the actual observed data and the predictions is minimized. This approach ensures that the model is well-fitted even for all misclassified observations.

In this regard, and with the aim of having an explicit and easily interpretable model, both a simple linear regression model and an ensemble training machine learning model were applied and their outcomes compared. In terms of simplicity, the term **"Tree-Fit"** is employed to characterize this particular approach. In the following sections, the steps taken in both the Simple Regression Methodology and Tree-Fit Modeling (referring the **fitting** process of **Ensemble Tree Learning Method**) are explained in further detail. The results and their interpretations will be analyzed in Section 3.3.

3.2.2 Methodology Phase 1: Grain Yield versus Climate Factors

Monthly average dew point and total precipitation data were grouped into a matrix corresponding to the three periods mentioned in Table 3.3. These scores were set as the independent variable to relate to that specific year's yield, making the dependent variable the annual yield for that specific grain in that year.

Since the harvest year for wheat and barley starts in September of the previous year and ends in the yield year, data for these grains spans 22 years. For corn, data spans 23 years, covering all selected provinces.

Then below steps are followed;

1. The number of observations for the independent variable is the yield of the selected provinces, which is set as the dependent variable “Y”.
2. The scores are determined from the matrices of “Total precipitation” and “dew point” observations, and these are grouped as "X" (see Section 3.2.2.1)
3. Total precipitation and dew point scores, X_{GP} are linked with yield Y, for each grain type G and province P as shown in below formula (also mentioned in Equation 3.1).

$$Y_{GP,year} = \beta_0 + \beta_1 * X_{GP1} + \beta_2 * X_{GP2} + ... + \beta_n * X_{GPn} + \epsilon$$

   where:

   - Y_{GP} is the yield for grain type G in province P.
   - X_{GP1}, X_{GP2}, ..., X_{GPn} are the scores derived from total precipitation and dew point for grain type G in province P.
   - β_0 is the intercept term.
   - β_1, β_2, ..., β_n are the coefficients corresponding to each independent variable.
   - ϵ is the error term.

4. The coefficients (β) are estimated through simple regression as below Equation 3.3, by applying a log transformation to Y to narrow down the confidence interval of the dependent variable. This ensures a more precise relationship between the climate factors and the yield. The details are given in Section 3.2.2.2.

$$Betas = regress(log(Y), [ones(22, 1), XGP1, XGP2, XGP3]) \tag{3.3}$$

5. For 5 provinces, annual yield for 3 grains are formulated, so 15 regression models were formed in the first place such as;

   - Konya_Wheat, Konya_Barley, Konya_Corn,
   - Yozgat_Wheat, Yozgat_Barley, Yozgat_Corn,
   - Polatlı_Wheat, Polatlı_Barley, Polatlı_Corn,
   - Adana_Wheat, Adana_Barley, Adana_Corn,
   - Urfa_Wheat, Urfa_Barley and Urfa_Corn.

3.2.2.1 Finding Scores Based on Dew Point and Total Precipitation Matrices

To find the scores, total precipitation (TP) and dew point (DP) are grouped based on periods aligned with the life cycle of the grain in the selected provinces. These periods are outlined in Table 3.3.

The following steps are taken:

- The related province's monthly total precipitation and monthly dew point figures are retrived from ECMWF. This data comes as a 2-size rectangular matrix, corresponding to observations in the related "$latitudeXlongitude$" points (in below examples they are respectively referred as "mXn") of the related province (larger the area, larger the matrix size) per month.

$$\textit{Jan 2000 values}: \begin{vmatrix} Lat1, Lon1 & Lat1, Lon2 & ... & Lat1, Lonn \\ Lat2, Lon1 & ... & ... & Lat2, Lonn \\ ... & ... & .... & \\ Latm, Lon1 & ... & ... & Latm, Lonn \end{vmatrix}$$

- The rectangular matrix of the related month is `reshaped` to form 1 column per month (a vector).

$$\textit{Jan 2000 values}: \begin{vmatrix} Lat1, Lon1 \\ Lat2, Lon1 \\ Lat3, Lon1 \\ ... \\ Latm, Lon1 \\ ... \\ Lat1, Lonn \\ ... \\ Latm, Lonn \end{vmatrix}$$

- Each vector-month within the same year is grouped in a rank from January to December, to form a $(latitude * longitude)$ X 12 size matrix.

$$\begin{vmatrix} Jan2000Lat1, Lon1 & Feb2000Lat1, Lon1 & ... & Dec2000Lat1, Lon1 \\ ... & ... & ... & ... \\ Jan2000Latm, Lonn & Feb2000Latm, Lonn & ... & Dec2000Latm, Lonn \end{vmatrix}$$

- **Dew Point** matrices are combined in period wise by taking **mean** of the corresponding month columns of the period as below MATLAB formula, resulting a vector resulting a vector [(latitude*longtitude) X 1] size.

$$DP_{Prov,Year,Period} = mean(Year(:, periodstartrank : periodendrank), 2)$$

- **Precipitation** matrices are combined in period wise by taking **sum** of the corresponding month columns of the period as below MATLAB formula, resulting a vector [(latitude*longtitude) X 1] size.

$$TP_{Prov,Year,Period} = sum(Year(:, periodstartrank : periodendrank), 2)$$

- Each of these $(latitude*longitude)$ size rows and 1 column matrices per period for years are transposed and compiled to make a matris of all years in that particular period per province making 22 years X $(latitude * longitude)$ size matrix. [2]

[2] For wheat the cycle starts in September of the previous year so 2022 is not concluded since it is related with 2023 yield results which would be out of scope, for corn on the other hand, years are between 2000 and 2022 making 23 years to work on, because sowing and harvest takes place at the same year.

$$MatrixDP_{Province,Period,allyears} = [DP'_{Province,Period,2021}; DP'_{Province,Period,2020}; DP'_{Province,Period,2019}; DP'_{Province,Period,2018}; DP'_{Provinceperiod,2017}; DP'_{Province,Period,2016}; DP'_{Province,Period,2015}; DP'_{Province,Period,2014}; DP'_{Province,Period,2013}; DP'_{Provinceperiod,2012}; DP'_{Province,Period,2011}; DP'_{Province,Period,2010}; DP'_{Province,Period,2009}; DP'_{Province,Period,2008}; DP'_{Province,Period,2007}; DP'_{Province,Period,2006}; DP'_{Province,Period,2005}; DP'_{Province,Period,2004}; DP'_{Province,Period,2003}; DP'_{Province,Period,2002}; DP'_{Province,Period,2001}; DP'_{Province,Period,2000};]$$

$$MatrixTP_{Province,Period,allyears} = [TP'_{Province,Period,2021}; TP'_{Province,Period,2020}; TP'_{Province,Period,2019}; TP'_{Province,Period,2018}; TP'_{Province,Period,2017}; TP'_{Province,Period,2016}; TP'_{Province,Period,2015}; TP'_{Province,Period,2014}; TP'_{Province,Period,2013}; TP'_{Province,Period,2012}; TP'_{Province,Period,2011}; TP'_{Province,Period,2010}; TP'_{Provinceperiod2009}; TP'_{Province,Period,2008}; TP'_{Province,Period,2007}; TP'_{Province,Period,2006}; TP'_{Province,Period,2005}; TP'_{Province,Period,2004}; TP'_{Province,Period,2003}; TP'_{Province,Period,2002}; TP'_{Province,Period,2001}; TP'_{Province,Period,2000};]$$

- Dew Point matrix and Precipitation matrix are compiled **per period** to make the ultimate matrix to work on it.

```
Weather_{Prov,Per,Gra}=[MatDP_{Prov,Per,Gra},MatTP_{Prov,Per,Gra}]
```

This structured approach ensures that the effects of precipitation and dew point on grain production are accurately captured, taking into account the essential growth stages in each province.

- The scores of the matrix "$Weather_{Province,Period,Grain}$" (PPG) are found via applying the below formulation in MATLAB to the ultimate matrix:

```
[coeffPPG,scorePPG,latentPPG,tsquared]= pca(zscore(PPG))
```

Thus, reducing the dimensionality of a large data set is accomplished through principal component analysis (PCA).

- To determine the number of scores to take into consideration, the following formula is applied:

```
latentPPG./sum(latentPPG)
```

- The result gives a 22 year-row (for Corn, 23 years) and 1 column matrix showing the relevancy percentage of the scores in descending order. The percentages are aggregated until it reaches at least $\%90$ and the number of scores making this sum is taken into consideration to put in regression formula (see Equation 3.3). The higher the relevancy, the number of scores is small and vice versa.

- There may be more than one score per period depending on the number of scores identified (see Equation 3.1).

- The above steps are applied for all periods and to be able to make predictions, scores of all periods are compiled in one term as below MATLAB formula explains:

$$X = [ones(22^3, 1), XGP1, XGP2, XGP3];$$

3.2.2.2 Finding Coefficients of the Scores

- For each period, the periods are combined, for Wheat and Barley to form below matrix structures for calculating the scores for each year and province, where each row represents a year and columns represent the compiled TP and DP for each period as below:

  Yield **2001**= Scores of TP and DP **2000**(Period 1); Scores of TP and DP 2001(Period 2) ; Scores of TP and DP 2001(Period 3)
  Yield 2002=Scores of TP and DP 2001(Period 1); Scores of TP and DP 2002(Period 2) ; Scores of TP and DP 2002(Period 3)
  ...
  Yield 2022= Scores of TP or DP 2021(Period 1); Scores of TP and DP 2022(Period 2) ; Scores of TP and DP 2022(Period 3)

- **For Corn** 23 years are studied since both sowing and harvest occur within the same calendar year. The structure for calculating the scores for corn is outlined below:

  Yield 2000=Scores of TP and DP 2000(Period 1); Scores of TP and DP 2000(Period 2) ; Scores of TP and DP 2000(Period 3)

---
[3] 23 years for Corn

Yield 2001= Scores of TP and DP 2001(Period 1); Scores of TP and DP 2001(Period 2) ; Scores of TP and DP 2001(Period 3)

...

Yield 2022=Scores of TP and DP 2022(Period 1); Scores of TP and DP 2022(Period 2) ; Scores of TP and DP 2022(Period 3)

- This structured approach ensures that the effects of precipitation and dew point on grain yield are accurately captured, considering the specific growth stages of wheat and barley in each province.
- The regression is applied and the β coefficients in equation are found as in Equation 3.3. These coefficients are only subject to making **estimations** [4] the results of which is shown under Section 3.3.1.2, in the 4th column of Table 3.9, Table 3.10 and Table 3.11.

3.2.2.3 Tree-fit (Ensemble Tree Learning Method) Prediction

On one hand, the series of Y by omitting the last figure $(1 : n - 1)$ is compared with the X starting from one year after Y until the end of the list $(2 : n)$. This comparison is conducted using the Tree-fit model with the following MATLAB formula:

```
Mdl=fitrensemble(X(2:n,:),
log(Y(1:n-1)));
```

Then the future Y figure (0) is predicted by utilizing the first observation of X with the following MATLAB formula:

```
f=predict(Mdl,X(1,:))
```

Given that the logarithm of Y is employed in the model, the outcome is derived by exponentiating f, yielding the prediction itself. This identical formula is then applied to the remaining observations in the same sequence as delineated above.

[4] for **predictions**, the formulation is explained under Section 3.2.2.4.

3.2.2.4 Simple Regression Prediction

For **prediction** in simple regression, akin to the prediction method in the tree-fit model, the series of Y is extracted from the first observation to the one before the last observation $(n-1)$, while X commences from one year before Y until the end of observations. The multipliers are determined using the following MATLAB formula:

```
[mm,~,~,~,~]= regress(Y(1:n-1),X(2:n,:))
```

The multipliers found from the previous year are applied to the equation of the current year, and the prediction for the future year is calculated using the following MATLAB formula:

```
Y_0= mm'*X(1,:)'
```

The same steps are repeated for the remaining observations in the same order, allowing future years to be predicted based on the actual or predicted figures from previous years.

3.2.3 Methodology Phase 2: Local Price versus Local Yield, Net Import and Global Prices

Panel data for three grains were collected, encompassing the total yield in the selected five provinces, entire country's net import figures accumulated over each period and global prices corresponding to the last month of each period which is selected among the criteria explained under Section 3.1.1.4. The equations were formulated according to the following criteria:

- The average monthly prices of the grains are computed based on the corresponding month's values, denominated in Turkish lira and adjusted for real prices (see also Section A). To accomplish this,
  - Local prices are sourced from the spot TMEX market, as well as local trading markets like Konya, Polatli, Yozgat, and Sanliurfa. These prices are reported per kilogram.

  - Global prices are obtained from CBOT future quotes for Wheat and Corn, while NMEX prices are used for Barley. The quotations are based on a bushel basis, with each bushel corresponding to 0.0272155 metric tons for Wheat, 0.021772 metric tons for Barley, and 0.0254 metric tons for Corn contracts. These prices are initially converted into kilograms and then into Turkish Lira.
  - Real TL prices are calculated by adjusting the current prices for the respective country's inflation rate.

- Annual import and export numbers are categorized into groups based on periods.

- Annual local yield numbers are derived by summing up the yields in the selected provinces (as they are the result of the Phase1 calculations) and divided into groups based on periods.

- The analysis is conducted for the years between 2010 and 2022, as local prices are not available before this timeframe.

The primary assumption underlying this model posits that the local price (LP) is influenced solely by three factors: local yield (Y), net imports (MX), representing the remaining yield produced inland and not exported, and the global price (GP) of the grain (G). Unlike the scoring process described in Section 3.2.2.1, actual numbers are directly inputted into the formula.

$$LP_{G,Per,year} = \beta_0 + \beta_1 * MX_{G,Per,year} + \beta_2 * Y_{G,Per,year} + \beta_3 * GP_{G,Per,year} + \epsilon$$

With 13 years under examination, there are 13 rows, and the dependent variables are grouped accordingly:

$$X_{Period} = [ones(13,1), MX_{Grain,Period}, LocalYield_{Grain,Period}, GP_{Grain,Period}]$$

for below models:

- Wheat_Per1, Wheat_Per2, Wheat_Per3, Wheat_Per4, Wheat_Per5

- Barley_Per1, Barley_Per2, Barley_Per3, Barley_Per4, Barley_Per5
- Corn_Per1, Corn_Per2, Corn_Per3, Corn_Per4, Corn_Per5

3.2.3.1 Tree-fit Prediction

On one hand, the series of Y representing the local price, by excluding the last figure $(1:n-1)$ is compared with X_{Period} starting from one year after Y until the end of the list $(2:n)$. This comparison is conducted using the tree-fit model via the following MATLAB formula:

```
Mdl=fitrensemble(X(2:n,:),log(LocalPrice(1:n-1)));
```

Then the future Y figure (0) is predicted is made by taking the first observation of X via below MATLAB formula.

```
f=predict(Mdl,X(1,:))
```

Since logarithm of Y is used in the model to narrow down the confidence interval of the dependent variable, the result is obtained by taking the exponential of f which represents the prediction. This process is repeated for the remaining observations in the same order as described above, allowing for the prediction of future year local prices using the actual/predicted figures from previous years.

3.2.3.2 Simple Regression Prediction

The beta multipliers are found through optimization using the simple regression formula in MATLAB, as shown below. In simple regression **predictions**, the series of Y is taken from first observation to the one before the last observation $(n-1)$ while X is starting from one year before Y until the end of observation.

```
[mm,~,~,~,~]= regress(LocalPrice(1:n-1),X(2:n,:))
```

The multipliers found from the previous year are applied to the equation of the current year, and the prediction for the future year is made using the following formula:

```
LocalPrice\_{0}= mm'*X(1,:)'
```

The same steps are repeated for the remaining observations in the dataset, following the same order.

3.3 Results and Interpretation

The research offers price prediction based on weather conditions, particularly total precipitation and dew point, utilizing a random forest decision tree fitting method, which constitutes an advanced machine learning model. These predictions can subsequently be integrated into decision tree frameworks alongside other decision layers. The price predictions exhibit consistency when employing the tree-fit model. Hence, barring substantial inflation or exchange rate fluctuations, the price predictions are anticipated to be reliable.

Within this framework, the predictions are structured and articulated through two sequential phases.

3.3.1 Phase 1_Prediction Results

3.3.1.1 Phase 1_Simple Regression Stats

As explained in Section 3.2.2.1, the number of scores are determined according to the consolidated effect of precipitation and dew point to the yield of the grain G in the particular province P. After that, optimisation is held via regression and some statistics are retrieved from the model with a unit root test applied as in Augmented Dickey Fuller (ADF). In this test, the null hypothesis is that Y follows a unit root process. When it fails to reject the null hypothesis it means the test may provide evidence that the yields in that particular grain and province pair is non-stationary. When it rejects the null hypothesis in favor of the alternative model it means the yield series is stationary. As seen in Table 3.5, when there are some failures in rejection in certain grain-province pair for actual yield series in "Unit Root Y" column, unit root is proven to be rejected in all grain-province pairs for the residuals between the actual input and the estimations as shown in "Unit Root R" column.

Table 3.5: Some Descriptive Statistics of Phase_1 Simple Regression Model

Provinces	**Grains**	**R2**	**F-stats**	**P-value**	**Unit Root Y**[5]	**Unit Root R**[6]
Konya	**Wheat**	0.66	2.60	0.06	failure to reject	reject
	Barley	0.54	1.59	0.22	failure to reject	reject
	Corn	0.61	1.55	0.24	reject	reject
Polatlı	**Wheat**	0.41	1.75	0.18	failure to reject	reject
	Barley	0.53	2.87	0.05	failure to reject	reject
	Corn	0.35	1.83	0.16	reject	reject
Yozgat	**Wheat**	0.4	1.07	0.44	failure to reject	reject
	Barley	0.57	2.18	0.1	failure to reject	reject
	Corn	0.37	0.7	0.71	reject	reject
Adana	**Wheat**	0.72	3.39	0.03	failure to reject	reject
	Barley	0.23	0.41	0.91	failure to reject	reject
	Corn	0.81	2.00	0.18	failure to reject	reject
Urfa	**Wheat**	0.14	0.41	0.86	failure to reject	reject
	Barley	0.33	1.22	0.35	failure to reject	reject
	Corn	0.58	1.37	0.31	failure to reject	reject

Following the optimization conducted using simple regression, the coefficients are found as shown in Table 3.6. The standard errors are also shown. The same scores are used in tree-fit model but coefficients are not generated as an output nor the standard errors since the ensemble method aggregates the errors of the base learner and builds the model by changing the data every time to fit and train data used by the base learner[37]. Tree-fit results are going to be discussed in Chapter Section 3.3.1.2.

Table 3.6: Wheat Model per Province_Coefficients and Standard Errors

Wheat Scores	**Konya W**	St E [7]	**Polatlı W**	St E	**Yozgat W**	St E	**Adana W**	St E	**Urfa W**	St E
Intercept	14.3691	0.0378	13.7896	0.0405	13.43734	0.0343	13.4868	0.0251	13.7904	0.0453
Per1_1st	- 0.0022	0.0017	-0.0066	0.0036	0.00025	0.0016	0.0022	0.0010	- 0.0015	0.0021
Per1_2nd	- 0.0035	0.0028	-0.0060	0.0067	- 0.00326	0.0031	0.0022	0.0015	0.0018	0.0041
Per1_3rd	- 0.0019	0.0080	-	-	- 0.0029	0.0119	- 0.0013	0.0029	-	-
Per1_4th	-	-	-	-	-	-	0.0015	0.0045	-	-
Per2_1st	0.0018	0.0019	0.0077	0.0040	0.00152	0.0020	- 0.0026	0.0011	0.0011	0.0022
Per2_2nd	0.0036	0.0025	0.0064	0.0046	0.00417	0.0021	0.0014	0.0015	0.0044	0.0037
Per2_3rd	- 0.0078	0.0062	-	-	-	-	-	-	-	-
Per3_1st	0.0052	0.0016	0.0055	0.0035	0.0013	0.0024	0.0003	0.0012	- 0.0004	0.0019
Per3_2nd	- 0.0050	0.0038	0.0042	0.0084	0.00025	0.0035	- 0.00005	0.0012	- 0.0023	0.0037
Per3_3rd	0.0035	0.0059	-	-	0.00066	0.0068	- 0.0103	0.0033	-	-
Error Variance	***0.0315***		***0.0362***		***0.02589***		***0.0139***		***0.0452***	

According to the number of scores and the error variances in Table 3.6, for wheat the least error variance belongs to Adana where one can conclude that total precipitation and dew point figures can best be related to the model however, Adana also needs 4 variables for Period 1 which is not the case for other provinces. On the other hand,

[6] Applied to yield outputs only

[6] Applied to residuals

[7] All standard errors display 90% confidence intervals for the coefficients

when looked at the results per province, the average accuracy is its best in Urfa and Polatlı respectively, where the number of scores are less than those of other provinces. It may be concluded that number of scores may also be an indicator of a strong relationship, the less they are the higher the correspondency.

Table 3.7: Barley Model per Province_Coefficients and Standard Errors

Barley Scores	**Konya B**	St E	**Polatli B**	St E	**Yozgat B**	St E	**Adana B**	St E	**Urfa B**	St E
Intercept	13.6832	0.0425	13.3143	0.0430	118.163	0.0637	97.260	0.0749	129.514	0.0993
Per1_1st	- 0.0019	0.0020	- 0.0089	0.0038	0.0012	0.0029	- 0.0011	0.0032	0.0061	0.0046
Per1_2nd	- 0.0046	0.0031	- 0.0160	0.0071	0.0040	0.0058	- 0.0040	0.0045	- 0.0062	0.0090
Per1_3rd	- 0.0105	0.0090	-		- 0.0472	0.0222	- 0.0067	0.0087	-	-
Per1_4th	-	-	-	-	-	-	- 0.0109	0.0134	-	-
Per2_1st	0.0013	0.0022	0.0082	0.0042	- 0.0026	0.0037	0.00002	0.0031	0.00001	0.0048
Per2_2nd	0.0039	0.0028	0.0103	0.0049	0.0033	0.0040	0.0009	0.0045	- 0.0119	0.0081
Per2_3rd	- 0.0040	0.0070	-	-	-	-	-	-	-	-
Per3_1st	0.0027	0.0018	0.0063	0.0038	- 0.0009	0.0044	- 0.0011	0.0036	0.0034	0.0041
Per3_2nd	0.0044	0.0043	0.0041	0.0089	0.0079	0.0065	0.0003	0.0035	0.0119	0.0081
Per3_3rd	0.0005	0.0066	-	-	- 0.0128	0.0126	- 0.0108	0.0099	-	-
Error Variance	***0.0398***	-	***0.0407***	-	***0.0893***	-	***0.1233***	-	***0.2169***	-

In Barley, according to stats table Table 3.7, the least error variance is in Konya, however the number of scores are less in Urfa and Polatlı again. Looking at the result comparison of prediction and actual yield, Polatli and Urfa have a balanced differences among yields even their average accuracy is not in best rank. So here, less number of scores may be a better reference for the model achievement.

Table 3.8: Corn Model per Province_ Coefficients and Standard Errors

Corn Scores	**Konya C**	St E	**Polatli C**	St E	**Yozgat C**	St E	**Adana C**	St E	**Urfa C**	St E
Intercept	12.5370	0.2054	8.7583	0.1861	4.9220	0.2302	13.5664	0.0516	12.6329	0.1581
Per1_1st	0.0016	0.0087	- 0.0131	0.0156	- 0.0158	0.0126	- 0.0069	0.0029	- 0.0107	0.0092
Per1_2nd	- 0.0168	0.0222	- 0.0504	0.0301	- 0.0051	0.0185	0.0017	0.0035	- 0.0025	0.0193
Per1_3rd	- 0.0366	0.0283	-	-	0.0015	0.0390	- 0.0018	0.0062	-	-
Per1_4th	0.0476	0.0444	-	-	- 0.0219	0.0548	0.0037	0.0074	-	-
Per1_5th	- 0.0173	0.0450	-	-	-	-	0.0181	0.0138	-	-
Per2_1st	0.0404	0.0106	0.0263	0.0149	- 0.0136	0.0112	- 0.0039	0.0028	0.0087	0.0099
Per2_2nd	0.0388	0.0266	-	-	- 0.0263	0.0351	- 0.0015	0.0052	- 0.0080	0.0158
Per2_3rd	-	-	-	-	- 0.0193	0.0462	- 0.0013	0.0064	- 0.0353	0.0346
Per2_4th	-	-	-	-	-	-	- 0.0019	0.0095	0.0253	0.0431
Per2_5th	-	-	-	-	-	-	- 0.0153	0.0135	-	-
Per3_1st	0.0310	0.0115	0.0140	0.0170	0.0091	0.0142	- 0.0041	0.0030	- 0.0272	0.0119
Per3_2nd	- 0.0097	0.0153	0.0125	0.0279	0.0153	0.0168	- 0.0007	0.0034	0.0038	0.0155
Per3_3rd	- 0.0213	0.0356	-	-	- 0.0028	0.0546	0.0098	0.0056	0.0116	0.0240
Per3_4th	0.0261	0.0405	-	-	-	-	0.0000	0.0074	0.0276	0.0408
Per3_5th	-	-	-	-	-	-	- 0.0193	-	0.0170	0.0373
Error Variance	***0.9706***	-	***0.7968***	-	***12.189***	-	***0.0613***	-	***0.5747***	**-**

Corn has a difference from other grains where 23 years are counted since the sowing and the harvest take place in the same year. Looking at the standard errors in Table 3.8 the best model would be assumed as Adana, which is true looking at the average of difference ratio of the yield and the prediction through years. Here Adana has the most score items which did not change its accuracy. On the other hand, Urfa has a high standard error but still have a balanced difference among years and not much deviating.

Under these statistical references, the results from both models for all grains are examined, where low accuracy is observed as detailed in Section 3.3.1.2. For this phase, the application of a simple regression model does not appear suitable for handling this layered type of data, which demands more components to achieve accurate predictions. In contrast, predictions from the tree-fit model continue to provide significantly closer results.

3.3.1.2 Phase 1_Result Comparison

After following the steps explained in Section 3.2.2.1, tree-fit and simple regression methods are applied for each province of 5, and each grain of 3, making 15 tables province and per grain. Under this section, for each grain, one different province table is put as an example for convenience. The whole comparison of all results in all provinces is going to be done in 4.

The last columns of the following three tables, are the "Simple Regression Estimation" results where the β coefficients are calculated via simple regression method using the input and output of **the same year**, in "Simple Regression Prediction" column on the other hand, the predictions are calculated with a one year lag between output and input as explained in Section 3.2.2.4.

1. The table for “Konya Wheat” prediction result comparison based on the model is shown in Table 3.9. Since the model takes a two year break between actual data and the prediction year, 2002 and 2001 results were not applicable.

   Some results shown in **bold** from the simple regression prediction are not coherent and produce implausible outcomes. On the other hand, as clearer comparison can be seen in Figure 3.2 tree-fit model predictions goes not much deviated from actual numbers (K_Wheat) when all years are taken into account. To see the accuracy better, simple regression **estimation** results are put instead

[8] Difference between the actual yield in the first column and the estimated value via simple regression in the previous column

Table 3.9: The Result Comparison Based on Models for Konya Wheat Yield

	Actual Yield	Tree-fit Pred	Sim Reg Pred	Simp Reg Est	Residuals[8]
2022	1,929,537	1,389,120	1,241,369	1,877,289	52,248
2021	1,579,839	1,663,092	1,913,623	1,593,344	-13,505
2020	1,920,700	2,312,098	2,129,507	1,913,701	6,999
2019	1,886,131	2,354,136	2,690,999	2,168,643	-282.512
2018	2,037,936	2,927,907	2,975,550	1,979,792	58,144
2017	2,192,410	2,038,244	2,148,729	2,064,221	128,189
2016	2,045,298	1,830,914	2,154,230	1,838,656	206,642
2015	2,554,256	1,662,156	2,480,171	2,338,901	215,355
2014	1,905,300	1,236,621	1,366,271	1,948,413	-43,113
2013	2,291,930	1,740,268	**527,590**	1,558,871	733,059
2012	1,570,660	1,319,875	**204,300,573**	1,424,981	145.679
2011	2,444,814	1,440,577	-25,974	2,271,139	173,676
2010	1,515,303	1,431,499	1,416,913	1,925,901	-410,598
2009	1,696,165	1,397,223	2,782,830	1,788,130	-91,965
2008	1,089,782	1,456,309	1,083,281	1,238,207	-148,425
2007	1,026,565	1,561,806	3,600,087	1,253,891	-227,326
2006	1,586,033	1,555,807	**4,789,915**	1,506,227	79,806
2005	1,343,977	1,633,592	**-755,745**	1,443,302	-99,325
2004	1,497,569	1,706,169	**-2,404,108**	1,434,392	63,177
2003	1,684,153	1,728,473		1,877,901	-193,748

of simple regression predictions due to the latter's high deviations. The reason of this failure in simple regression **prediction** model can be interpreted as the poor statistical significance of the simple regression model which is discussed in the previous section, Section 3.3.1.1.

2. The table for "Yozgat Barley" prediction result comparison based on the model is as shown in Table 3.10. Since the model takes a two year break between actual data and the prediction year, 2002 and 2001 results were not applicable.

   As it can be seen, some results shown in **bold** of simple regression prediction gives way far results than the actual ones. Tree fit seems a better coherent fit. The comparison of the tree-fit model predictions and the actual numbers (Y_Barley) shown in Figure 3.3 also shows a coherency between them for all years, where only simple regression estimation results are put instead of simple regression **predictions** to make a better comparison.

3. The table for "Adana Corn" prediction result comparison based on the model is as shown in Table 3.11. Since the model takes a two year break between actual

[9]Difference between the actual yield in the first column and the estimated value via simple regression in the previous column

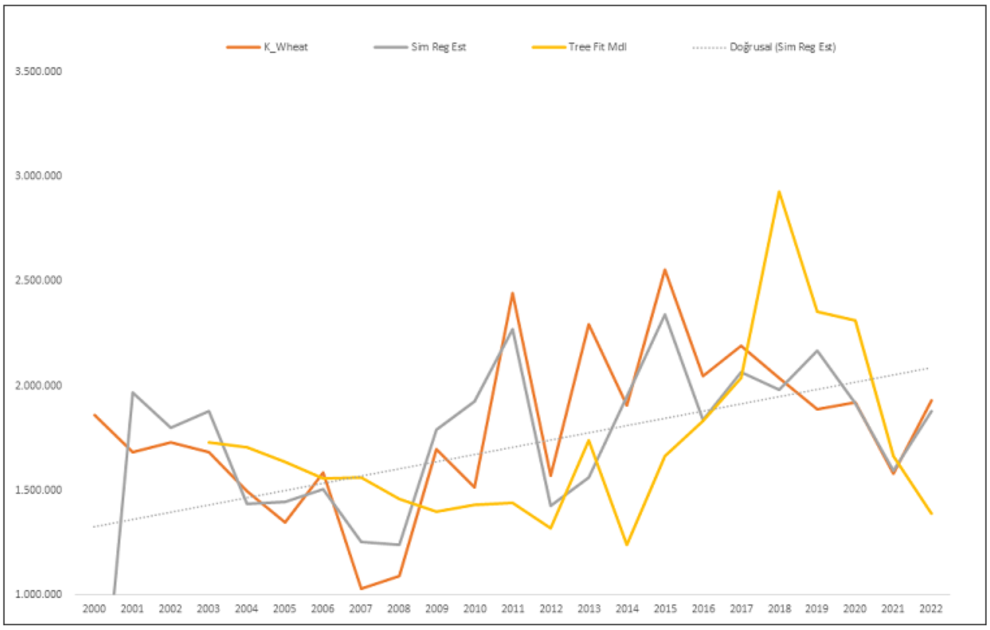

Figure 3.2: Konya Wheat Yield Comparison of St Reg Estimation and Tree-fit Model Prediction

Table 3.10: The Result Comparison of Models for Yozgat Barley Yield

	Actual Yield	Tree-fit Pred	Sim Reg Pred	Simp Reg Est	Residuals[9]
2022	158,294	165,119	101,511	143,392	14,902
2021	108,417	174,795	163,476	123,361	-14,944
2020	216,085	77,570	148,304	151,414	64,671
2019	103,813	92,190	**-33,644**	131,845	-28,032
2018	87,062	100,038	169,727	72,890	14,172
2017	84,341	60,536	71,173	102,326	-17,985
2016	85,120	130,633	83,590	80,125	4,996
2015	111,34	123,927	21,479	115,719	-4,379
2014	95,150	174,313	317,950	101,815	-6,665
2013	97,552	153,785	212,670	137,954	-40,402
2012	93,372	103,603	336,659	136,174	-42,802
2011	108,696	182,407	739,642	134,096	-25,400
2010	124,863	191,257	185,583	147,772	-22,909
2009	179,282	193,032	**-1,246,078**	205,281	-25,999
2008	158,392	199,501	**-480,869**	174,716	-16,324
2007	193,354	200,754	169,008	129,540	63,814
2006	247,959	190,430	**-268,764**	186,460	61,499
2005	250,241	173,858	154,111	167,032	83,209
2004	223,707	153,268	**-126,097**	198,282	25,425
2003	151,410	155,149		206,870	-55,460

data and the prediction year, 2001 and 2000 results were not applicable.

[10]Difference between the actual yield in the first column and the estimated value via simple regression in the previous column

Figure 3.3: Yozgat Barley Yield Comparison of St Reg Estimation and Tree-fit Model Prediction

Table 3.11: The Result Comparison of Models for Adana Corn Yield

	Actual Yield	Tree-fit Pred	Sim Reg Pred	Simp Reg Est	Residuals[10]
2022	888,348	822,864	1,229,516	928,596	- 40,248
2021	810,145	887,606	1,045,660	801,292	8,853
2020	819,978	854,277	544,352	881,650	- 61,.672
2019	717,802	1,033,197	799,587	627,457	90,345
2018	842,697	628,330	576,350	840,502	2,195
2017	1,036,130	862,764	**-379,736**	904,528	131,602
2016	1,086,606	1,329,161	**5,404,186**	1,113,611	- 27,005
2015	1,015,428	1,468,258	2,012,186	1,046,717	- 31,289
2014	1,005,651	833,396	**-997,068**	1,064,188	- 58,537
2013	915,284	506,583	**-272,331**	893,485	21,799
2012	682,462	706,580	2,759,473	800,118	- 117,656
2011	760,744	1,057,752	**-2,849,139**	793,897	- 33,153
2010	748,160	722,244	110,668	733,431	14,729
2009	894,099	703,228	**-182,930**	910,203	- 16,104
2008	1,036,626	665,305	1,219,935	958,328	78,298
2007	1,013,099	620,272	**-1,365,226**	885,070	128,029
2006	1,014,235	562,174	315,536	894,207	120,028
2005	1,014,668	485,018	**-244,589**	862,245	152,423
2004	722,394	424,703	**-166,326**	878,023	- 155,629
2003	506,575	388,872	319,443	464,500	42,075
2002	379,931	398,023		429,580	- 49,649

In the table, the results shown in **bold** of simple regression prediction does not give coherent results as tree-fit model did, also the prediction do not make any sense. The comparison of the tree-fit model predictions and the actual numbers (A_Corn) shown in Figure 3.4 also shows a coherency between them for all years, where only simple regression estimation results are put instead of simple regression predictions due to the latter's high deviations. The reason is again interpreted as from the poor statistical significance of the simple regression model which will be discussed in the following section.

Figure 3.4: Adana Corn Yield Comparison of St Reg Estimation and Tree-fit Model Prediction

3.3.2 Phase 2_Prediction Results

3.3.2.1 Phase 2_Simple Regression Stats

As explained in Section 3.2.3, X is made of three factors (total yield of 5 provinces as local yield, Turkey's net import and global price in international future markets) and one intersection point to conclude the dependent varable local price. These models are formed in 5 period-basis each period is determined by taking the future contract value dates in international markets as explained under Table 3.4. There is no score preparing, instead actual numbers are put directly in the model. Optimisation is done via regression and the following coefficients are found as shown in Table 3.12 with the corresponding standard errors. The accuracy of the results are going to be discussed in Section 3.3.2.2.

The tests done in section, together with the close prediction results may be interpreted as the powerful data and adequacy of the X variables to predict Y values. As clearly seen in Table 3.12, the standard errors are negligible so the accuracy of the coefficients are expected to be high for all periods. The prediction results also confirms this high accurancy, both in simple regression and tree-fit is giving high accurancy as it is going to be shown in Section 3.3.2.2.

Table 3.12: Wheat Model per Period_ Coefficients and Standard Errors

Wheat Inputs	**Period 1**	St E [11]	**Period 2**	St E	**Period 3**	St E	**Period 4**	St E	**Period 5**	St E
Intercept	0.1238	0.3193	0.0554	0.3368	0.0098	0.4530	0.3383	0,4794	0.5237	0.3690
Local Yield	-7.32E-11	5.71E-07	-1.32E-10	9.09E-07	-4.27E-10	1.25E-06	-2.39E-10	1,32E-07	-2.00E-10	6.89E-08
Net Import	2.06E-07	0.0021	3.74E-07	0.0034	3.85E-07	0.0045	2.59E-08	0,0048	-7.52E-08	0.0025
Global Price	0.2513	0.0119	0.2664	0.0116	0.4559	0.0209	0.3453	0,0194	0.3624	0.0159
Error variance	***0.0035***		***0.0096***		***0.0420***		***0.0286***		***0.0115***	

In Barley, according to stats table Table 3.13, local price has the lowest coefficient accompanied by low standard error showing that local yield has a high correlation with the local price, next is net import figures for all periods.

Table 3.13: Barley Model per Period_Coefficients and Standard Errors

Barley Inputs	**Period 1**	St E	**Period 2**	St E	**Period 3**	St E	**Period 4**	St E	**Period 5**	St E
Intercept	0.3152	0.2825	0.5022	0.3126	0.6249	0.2595	0.5342	0.2787	0.5530	0.3012
Local Yield	-4.33E-10	3.09E-06	-7.04E-11	4.95E-06	3.09E-10	4.04E-06	-2.67E-10	4.57E-06	4.77E-10	4.00E-07
Net Import	-2.90E-07	0.0055	-9.35E-07	0.0089	-1.36E-06	0.0007	-9.96E-07	0.0079	-6.57E-07	0.0057
Global Price	0.7618	0.0550	0.6799	0.0357	0.6410	0.0271	0.6211	0.0303	0.5291	0.0354
Error variance	***0.0210***		***0.0195***		***0.0222***		***0.0310***		***0.0280***	

Corn has the same result of variances which can be considered as low, the lowest coefficients are again belonging to local yield as can be seen in Table 3.14. For each grain an example is going to be shown in graphic forms and the comparison of each model can be seen in a random period in the next section per grain, it confirms the accuracy.

Table 3.14: Corn Model per Period_ Coefficients and Standard Errors

Corn Inputs	**Period 1**	St E	**Period 2**	St E	**Period 3**	St E	**Period 4**	St E	**Period 5**	St E
Intercept	0.5949	0.1206	0.4986	0.1800	0.4251	0.1768	0.4505	0.3291	0.4451	0.2372
Local Yield	-1.78E-10	1.51E-10	-1.81E-10	3.34E-10	-6.63E-10	3.32E-10	2.61E-10	6.24E-10	-3.31E-10	2.84E-10
Net Import	-7.94E-07	3.07E-07	-9.88E-07	6.80E-07	-5.34E-07	6.68E-07	-9.94E-07	1.24E-06	-4.11E-07	6.12E-07
Global Price	0.3607	0.0209	0.3889	0.0263	0.4533	0.0285	0.3703	0.0489	0.3233	0.0336
Error variance	***0.0191***		***0.0413***		***0.0324***		***0.0245***		***0.0167***	

[11] All standard errors display 90% confidence intervals for the coefficients

The other tests are done for all models, the results of which is given in Table 3.15. As it can be seen, the R^2 results, which show the goodness of fit, are very close to 0.99 level, so would be counted as highly acceptable. Also, the $Fstatistic$ is higher than the critical value, meaning the difference among groups is deemed statistically significant. P values need to be looked at where they are all 0, showing the model would be significant, in other words X variables do help predict Y. This statement would be confirmed by the close prediction achievements in simple regression model as run in Section 3.3.2.2. Also unit root existence are all rejected with $p\ values$ as 0.001 which confirms the clearness of the unit root test result.

Table 3.15: Some Descriptive Statistics of Phase_2 Simple Regr Model

Periods	**Grains**	**R2**	**F-stats**	P value	**Unit Root**
Period 1	Wheat	0.9951	608.82	0	reject
	Barley	0.967	87.83	0	reject
	Corn	0.9664	86.2	0	reject
Period 2	Wheat	0.9901	300.21	0	reject
	Barley	0.9869	226.39	0	reject
	Corn	0.9607	73.36	0	reject
Period 3	Wheat	0.9731	108.68	0	reject
	Barley	0.986	211.87	0	reject
	Corn	0.9734	109.73	0	reject
Period 4	Wheat	0.9766	125.39	0	reject
	Barley	0.9782	134.81	0	reject
	Corn	0.9747	115.75	0	reject
Period 5	Wheat	0.9914	344.04	0	reject
	Barley	0.9813	157.8	0	reject
	Corn	0.9786	136.99	0	reject

This statement would be confirmed by the high prediction achievements in simple regression model. As it is going to be seen better in next section, Section 3.3.2.2, the predictions on tree-fit model are also showing steadiness on giving close values. As understood from the outcomes and the interpretations, for all grains, a coherency between the predictions and the results are applicable for all periods.

3.3.2.2 Phase 2_ Result Comparison

In Phase 1 the model was build on provinces since the weather conditions were belonging to provinces. In Phase 2, the model is based on periods rather than provinces, since the yield corresponds to total production of all 5 provinces. Local Price and Net Import also imply the figures in the whole country and global price indicates the international market prices so provinces are not related in this phase. As it can be seen in Section 3.2.3 the equation form are applied in period wise. Also in this phase no score grouping is taking place, instead the actual figures are put in the formula to get X where Y is Price. Again tree-fit and simple regression methods are applied for each period of 5, and each grain of 3, making 15 tables per period and per grain. Unlike Phase 1, there are 13 years to analyze in Phase 2.

Under this section, for each grain, one different period outcome is put as an example for convenience. The whole comparison of all results in all periods is going to be done in Chapter 4.

1. For wheat local price, Period 1 result comparison is shown as an example in Table 3.16. Simple Regression estimation is also demonstrated to confirm the accuracy of the coefficients. Since the model takes a two year break between actual data and the prediction year, 2011 and 2010 results were not applicable.

Table 3.16: Result Comparison of Models for Wheat Price in Period 1

Years	**Actual W Price**	**Tree-fit Pred**	**Sim Reg Pred**	**Sim Reg Est**
2022	3.2247	0.7094	1.7094	3.2133
2021	1.0532	0.5814	0.9635	1.0810
2020	0.7778	0.5915	0.6653	0.8436
2019	0.6826	0.5810	0.6701	0.6245
2018	0.5466	0.5861	0.6896	0.6193
2017	0.5910	0.5853	0.6282	0.6025
2016	0.5836	0.5857	0.5790	0.5630
2015	0.6533	0.5699	0.6459	0.5985
2014	0.6358	0.5495	0.6965	0.5405
2013	0.5940	0.5284	0.4252	0.5962
2012	0.4690	0.5954		0.5328

There are two important things in this table requires attention. One is simple regression prediction gives closer predictions than that of Phase 1, the second one is, even the inflation is applied to the prices, the volatility in the last figure of data set is high. This may be due to the asymmetric information that the price setters could not adjust the price level to the high and very fast increasing inflation rate, also the official announced inflation rate difference to the real life inflation. This close accuracy also can be seen in the comparison in the

Figure 3.5. Tree fit is following a balanced path however simple regression prediction model gives better coherence.

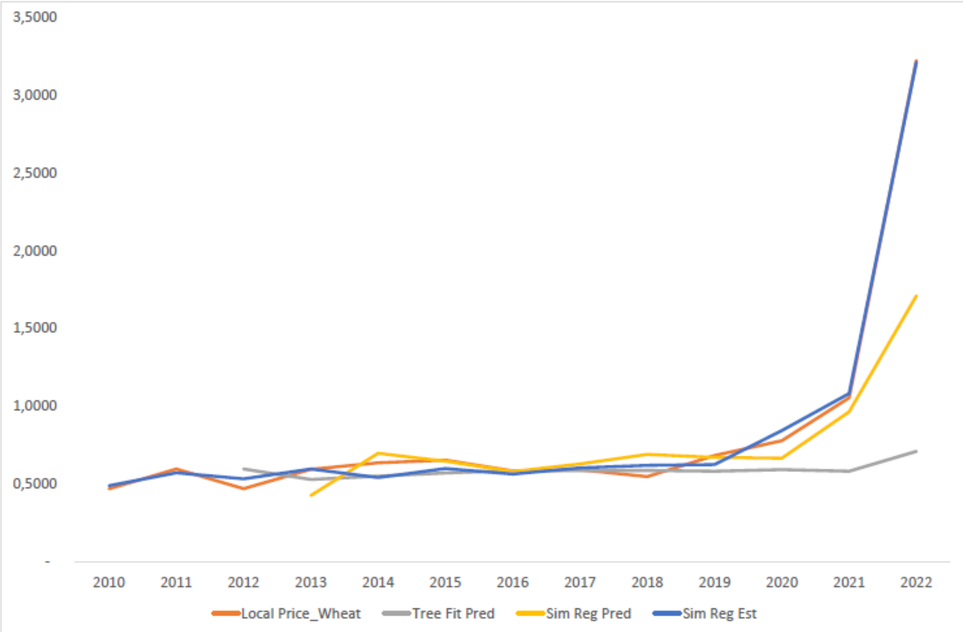

Figure 3.5: Period1_Local Wheat Price Comparison of Tree-fit and St Reg Predictions versus St Reg Estimation

2. For Barley, Period 3 is taken as an example, in this period price prediction result comparison can be seen in Table 3.17. As simple regression estimation confirms the accuracy of the coefficients, so the simple regression prediction can be applied with serenity. The model takes a two year break between actual data and the prediction year, therefore 2011 and 2010 results are not applicable so cannot be shown.

   The numbers are going align compared to the prediction results in Phase 1 inputs. Only in 2022 where the inflation rate made a jump, there has been a huge difference in Barley price probably the models could not adjust that drastically. The history of the data in tree-fit model may be the reason for not quickly adjusting in the model, however as can be seen cleare in below figure, Figure 3.6 tree-fit is still performing stable while simple regression has closer predictions than tree-fit.

Table 3.17: Result Comparison Based on Models for Barley Price in Period 3

Years	Actual Price	Tree-fit Pred	Sim Reg Pred	Sim Reg Est
2022	4.3822	0.6939	1.2104	4.3343
2021	1.6286	0.6120	0.7129	1.5862
2020	0.7511	0.5097	0.7505	0.7188
2019	0.6304	0.4963	0.5661	0.9221
2018	0.4942	0.4966	0.6228	0.7402
2017	0.5872	0.4829	0.4328	0.5473
2016	0.5110	0.4775	0.5759	0.5363
2015	0.4227	0.4922	0.5238	0.3091
2014	0.5543	0.4731	0.5275	0.5015
2013	0.4782	0.4706	0.4448	0.3175
2012	0.5019	0.4412		0.4307

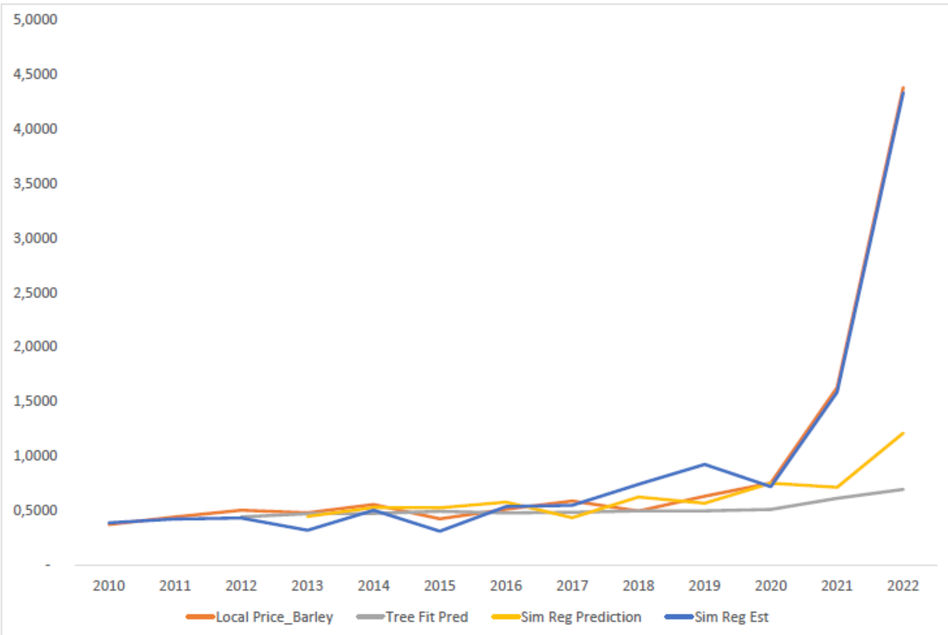

Figure 3.6: Period3_Local Barley Price Comparison of Tree-fit and St Reg Predictions versus St Reg Estimation

3. Lastly for Corn the example is of Period 5, the price prediction results can be compared via below table Table 3.18.

4. For Corn prices, the same prediction pattern is followed only there has been a huge deviation in simple regression prediction in 2022 as can be followed from Figure 3.7, other than that simple regression prediction is going align but one cannot exactly conclude on the better model for example treefit model predic-

Table 3.18: Result Comparison Based on Models for Corn Price in Period 5

Years	Actual C Price	Tree-fit Pred	Sim Reg Pred	Sim Reg Est
2022	2.7415	0.8126	5.0564	2.8416
2021	2.2084	0.4777	0.5632	1.9809
2020	0.6812	0.4398	0.4267	0.8058
2019	0.4489	0.4387	0.4247	0.4423
2018	0.3983	0.4448	0.5116	0.5411
2017	0.4136	0.4502	0.5685	0.3921
2016	0.4564	0.4489	0.6841	0.4663
2015	0.3936	0.4639	0.4464	0.3516
2014	0.4607	0.4650	0.4704	0.3412
2013	0.4805	0.4575	0.4500	0.3388
2012	0.4701	0.4452		0.5293

tion is performing better in 2018 and in 2020 simple regression is performing better. The model selection would depend on the preference of the researcher.

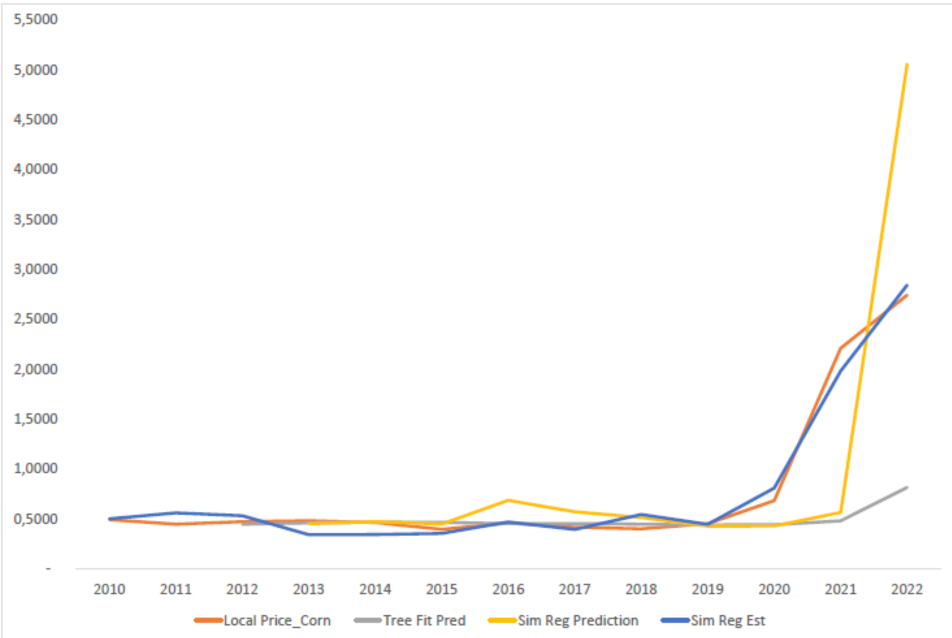

Figure 3.7: Period 5_Local Corn Price Comparison of Tree-fit and St Reg Predictions versus St Reg Estimation

3.3.3 Developing a Future Price Index

In this study, the objective is to establish a connection between climate conditions and the prices of specific agricultural products. This is pursued in two stages: first, by attempting to use climate data to predict yield figures, and then by linking local yield to price predictions along with other factors such as net import and global price. Ultimately, the goal is to demonstrate how climate can influence prices through this connection. Other factors assumed to be constant include freight, import taxes, exchange rate fluctuations, and inflation.

The final phase of the study involves the creation of an index. Initially, various price index examples are explored. Then, the outcomes of the relationship or the elements in the models studied in this paper are revisited to propose a future price index form. In local terms, the TMEX Grain Index is examined. TMEX, the Turkish Mercantile Exchange, is a nationwide exchange where Electronic Warehouse Receipts (EWR) representing agricultural products stored in licensed warehouses are bought and sold in a dedicated electronic platform by investors, producers (farmers), and merchants.

TMEX has developed indices on grain prices, both in single product terms (simple index) and in certain groups of them (composite index), quoted in the TMEX EWR Market. Composite commodity indices include the Grain Index (comprising Wheat, Barley, and Corn), Wheat Index, Barley Index, and Corn Index, while simple indices include the Bread Wheat Index, Durum Wheat Index, Grade 1 Corn Index, and Grade 2 Corn Index. As indicated on their official website [46], the components are selected based on specific criteria:

- The grain type produced and imported must be above 1 million tons on average over the last three years.
- They must be traded in the TMEX market above 100 thousand tons in volume on average over the last three years, with an average daily volume above 500 tons during the same period.
- They must be traded on at least 60% of business days in a year.

Other indices prepared by international entities include the International Grain Council (IGC) Grains and Oilseeds Index (GOI)[19] and the MSCI (Morgan Stanley Capital International) Equity Indices [28]. Similar to TMEX, these indices aim to measure the market performance of selected commodities or equities over a specific period.

The basic principles of price indices can be listed as follows:

- A price index level is calculated for a particular date, usually taken as the base level of 1000 when proportioning to future levels.
- The price index level is not the exact price of a particular commodity but an adjusted market level including the weights of selected commodities meeting specific criteria (such as those mentioned in the TMEX index).
- Other criteria for measuring the index include the number of trading days or the proportion of trading on a particular day.

These principles indicate that price indices essentially show the price performance of actual price data, whether from the spot or futures market. In this study, however, a price prediction model is developed, and future price announcements can be made based on these predictions. It will be up to the preference and decisions of future index structurers to select the weights of the commodities based on any criteria they deem appropriate, taking into account the future prices predicted by this tree-fit model.

CHAPTER 4

CONCLUSION

4.1 Results Accuracy

4.1.1 Phase 1 Result Accuracy

In this study, a substantial amount of data has been analyzed to establish reasonable correlations. For better clarity, the data points examined can be summarized as follows:

- 546 data points for total precipitation and 546 for dew point, totaling 1092 data points per month in Konya. Over 12 months, this amounts to 13,104 data points per year, and over 23 years, it makes 301,392 weather forecasts..
- 121 data points for total precipitation and 121 for dew point, totaling 242 data points per month in Polatlı. Over 12 months, this equals 2,904 data points per year, and over 23 years, it makes 66,792 weather forecasts.
- 441 data points for total precipitation and 441 for dew point, totaling 882 data points per month in Yozgat. Over 12 months, this results in 10,584 data points per year, and over 23 years, it makes 243,432 weather forecasts.
- 651 data points for total precipitation and 651 for dew point, totaling 1302 data points per month in Adana. Over 12 months, this yields 15,624 data points per year, and over 23 years, it makes 359,352 weather forecasts.
- 441 data points for total precipitation and 441 for dew point, totaling 882 data points per month in Şanlıurfa. Over 12 months, this amounts to 10,584 data

points per year, and over 23 years, it makes 243,432 weather forecasts.

Thus, considering weather forecasts only, over 1,000,000 data points have been attempted to be linked with total yield for Phase 1. This extensive dataset has been initially grouped according to the sowing, earing, and harvest periods, and then scores have been retrieved. The relationship has been established via simple regression and tree-fit modeling approaches, as described in detail in Section 3.2.2.1, under Section 3.2.2.

The results are discussed in Section 3.3.1.2, where it is observed that tree-fit modeling performs better, as it provides smoother predictions without drastic changes compared to simple regression. However, it is noted that the data correlations found to be not strong as desired, as indicated in Section 3.3.1.1. This is understandable because total precipitation and dew point alone may not be sufficient to predict yield accurately, as yield may be influenced by various factors beyond weather conditions, such as grain prices and profit expectations.

Moreover, the tree-fit predictions appear to be consistent with the actual results for all grain types and provinces, as evidenced by the comparison provided in Table 4.1, Table 4.2 and Table 4.3. The percentage difference between the predictions and the actual figures (Δ) is presented for each grain type, highlighting the effectiveness of the tree-fit modeling approach. Simple regression predictions, on the other hand, yielded unrelated and inconsistent results and are therefore excluded from the demonstration.

When examining Table 4.1, the comparison table for prediction precision, the tree-fit model appears to perform best in Konya compared to other provinces, with the average precision between the selected periods being almost zero. However, the most outstanding individual performance is observed in Polatlı, indicated by the numbers in *italic*, which represent the highest accuracy for that year among provinces. This superior accuracy in Polatlı could be attributed to its smaller area, which may result in more uniformly distributed data. On the other hand, the large yield of wheat in Konya, as indicated in Section A, Table A.2 may be the reason why the outcome is better performing than other provinces. Conversely, the worst accuracy is observed in Urfa for the year 2019, where the actual yield exhibits sharp fluctuations, as seen in Table A.2, the tree-fit model could not accurately predict these fluctuations, which

Table 4.1: Tree-fit Model Yield Prediction Preciseness_Wheat

Years	**∆_Konya**	**∆_Polatli**	**∆_Yozgat**	**∆_Adana**	**∆_Urfa**
2022	-28%	*-10%*	-11%	50%	19%
2021	5%	29%	136%	*4%*	-36%
2020	20%	19%	*6%*	8%	-54%
2019	25%	*5%*	28%	10%	108%
2018	44%	*-4%*	30%	16%	22%
2017	-7%	*2%*	5%	23%	13%
2016	-10%	*-7%*	11%	53%	-23%
2015	-35%	18%	*-4%*	21%	-17%
2014	-35%	36%	*14%*	22%	28%
2013	-24%	-60%	-10%	23%	*1%*
2012	-16%	*-14%*	*-14%*	*14%*	-18%
2011	-41%	-17%	-25%	*8%*	-12%
2010	6%	*1%*	-15%	18%	-7%
2009	-18%	-21%	-19%	*6%*	-27%
2008	34%	20%	10%	*-8%*	50%
2007	52%	95%	18%	*-10%*	11%
2006	*-2%*	-9%	6%	-6%	*2%*
2005	22%	-19%	-19%	-26%	*-7%*
2004	14%	-8%	*1%*	-16%	9%
Avg	***-0.33%***	***2.91%***	***7.71%***	***11.03%***	***3.34%***

is understandable given the inherent variability in yield data.

In Table 4.2 the comparison table for prediction precision, Konya exhibits the highest average accuracy among the provinces. Additionally, Konya and Yozgat demonstrate the best individual prediction performances, with five instances of top accuracy marked in *italic*. Conversely, Urfa shows the poorest predictions, which could be attributed to the fluctuating numbers in actual yield data. Similar fluctuations in actual yield are observed in other provinces such as Konya, Polatlı, and Adana for the years 2007-2008, where the actual numbers nearly halved compared to other years (see Table A.3). These sharp fluctuations pose challenges for the prediction model to align with, thus impacting accuracy.

In Table 4.3 corn appears to be the most challenging grain to predict, as indicated by its lower accuracy ratio compared to other grains. Surprisingly, Urfa demonstrates the best average performance in predicting corn yield, while Yozgat shows the highest accuracy in individual predictions. However, it's worth noting that Yozgat has the lowest corn yield among the provinces, suggesting that lower production figures may

Table 4.2: Tree-fit Model Yield Prediction Preciseness_Barley

Years	**Δ_Konya**	**Δ_Polatli**	**Δ_Yozgat**	**Δ_Adana**	**Δ_Urfa**
2022	-17%	*-5%*	*4%*	77%	171%
2021	-14%	*12%*	61%	60%	286%
2020	-35%	-39%	-64%	*2%*	55%
2019	-15%	-22%	-11%	*6%*	13%
2018	-8%	*-4%*	15%	36%	45%
2017	9%	29%	-28%	63%	*8%*
2016	*-1%*	6%	53%	25%	170%
2015	-21%	27%	*11%*	35%	-29%
2014	*-5%*	55%	83%	54%	66%
2013	*-21%*	-65%	58%	54%	56%
2012	-28%	14%	*11%*	37%	-26%
2011	*-13%*	*-13%*	68%	14%	23%
2010	27%	-17%	53%	-31%	*16%*
2009	*-1%*	-16%	8%	-25%	-4%
2008	74%	46%	*26%*	77%	106%
2007	60%	146%	*4%*	60%	-20%
2006	-17%	-7%	-23%	*2%*	-20%
2005	9%	-15%	-31%	*6%*	-31%
2004	3%	*1%*	-31%	36%	-15%
Avg	***-0.72%***	***7.02%***	***14.04%***	***19.63%***	***45.81%***

lead to better predictions.

Corn production requires substantial amounts of water, which may not be readily available in some provinces like Adana. Consequently, not all years may exhibit a smooth trend in corn production due to water availability issues. This volatility stemming from varying water availability could be a contributing factor to the deviated predictions observed.

Overall, the test results of inadequacy would also be understandable because only total precipitation and the dew point would not be the only effect on yield. The yield may have a random walk behavior and obviously not subject to seasonality because farmers take many other effects into consideration while choosing the grain to produce, such as the price of the previous year so the expectations of the profit to be earned [30]. Despite the challenges, the tree-fit model appears to provide reasonable predictions, albeit not as efficient as desired. However, it still performs better than simple regression. Therefore, it can be concluded that in scenarios with limited variables, the tree-fit model could be utilized to provide relatively significant results,

Table 4.3: Tree-fit Model Yield Prediction Preciseness_Corn

Years	∆_Konya	∆_Polatli	∆_Yozgat	∆_Adana	∆_Urfa
2022	-96%	-56%	1312%	*-7%*	-17%
2021	-92%	-98%	-97%	*10%*	-23%
2020	-67%	-44%	-84%	*4%*	-81%
2019	-44%	-91%	-60%	44%	*4%*
2018	-68%	-70%	62%	*-25%*	87%
2017	-87%	-40%	48%	*-17%*	27%
2016	-77%	-37%	33%	*22%*	65%
2015	-90%	-51%	*-8%*	45%	-79%
2014	-95%	-58%	140%	*-17%*	-73%
2013	-82%	-34%	-48%	-45%	*-4%*
2012	-75%	54%	-43%	*4%*	-81%
2011	*-14%*	58%	-49%	39%	-80%
2010	*0%*	23%	358%	-3%	-69%
2009	*-1%*	11%	56%	-21%	-62%
2008	33%	*0%*	33%	-36%	-36%
2007	104%	36%	25%	-39%	*3%*
2006	171%	18%	*0%*	-45%	44%
2005	106%	*5%*	-52%	-52%	32%
2004	170%	112%	*-34%*	-41%	278%
Avg	***-15.95%***	***13.64%***	***83.79%***	***-9.55%***	***-3,38%***

since its predictions seems to be coherent with the actual results in all grain types and provinces. This conclusion is supported by the comparison tables in this section, which demonstrate the performance of the tree-fit model in predicting yield compared to actual results.

4.1.2 Phase 2 Result Accuracy

Considering the volatility in grain prices driven by the pandemic and the Ukrainian war, as well as the discrepancy between actual and announced inflation rates, the fluctuation in grain prices remained a challenge even when using real price data in the study. Despite these challenges, the predictions for 2021 and 2022 appear to be consistent with the actual results for all grain types and periods, both in the Tree-fit and Simple Regression Predictions. Detailed comparison tables provide insights into the performance of the predictions relative to actual results Table 4.4 , Table 4.5 and Table 4.6 per ***period*** as the symbol ∆ represents the percentage difference between the predictions and the actual figures (TF:Tree-fit Model; SR: Simple Regression Model).

Table 4.4: Price Prediction Preciseness Between Models_Wheat

Years	ΔTFW1	ΔSR1	ΔTF2	ΔSR2	ΔTF3	ΔSR3	ΔTF4	ΔSR4	ΔTF5	ΔSR5
2022	-78%	-47%	-80%	-44%	-85%	-67%	-82%	-63%	-75%	73%
2021	-45%	-9%	-49%	-18%	-55%	-35%	-57%	-38%	-75%	-54%
2020	-24%	-14%	-27%	-8%	-35%	-35%	-40%	-43%	-46%	-20%
2019	-15%	-2%	-25%	-9%	-16%	-34%	-18%	-47%	-28%	-28%
2018	7%	26%	14%	32%	17%	16%	6%	0%	0%	9%
2017	-1%	6%	-7%	6%	-2%	-3%	6%	7%	4%	6%
2016	0%	-1%	-2%	-4%	1%	12%	-1%	3%	-3%	12%
2015	-13%	-1%	-8%	6%	12%	-110%	8%	-85%	-2%	-32%
2014	-14%	10%	-11%	-22%	-22%	-15%	-15%	-8%	-17%	-10%
2013	-11%	-28%	-12%	-27%	1%	-12%	10%	-6%	-6%	-13%
Avg	***-19%***	***-6%***	***-21%***	***-9%***	***-18%***	***-28%***	***-18%***	***-28%***	***-25%***	***-6%***

In Table 4.4 the predictions under the simple regression method appear to yield better results in average terms for all periods except for Period 3 and Period 4. However, when considering individual predictions, the 2015 Period 3 and 4 predictions show a significant error in the simple regression model, resulting in a failure in average terms. On the other hand, the tree-fit method maintains a steady prediction trend.

Table 4.5: Price Prediction Preciseness Between Models_Barley

Years	ΔTFW1	ΔSR1	ΔTF2	ΔSR2	ΔTF3	ΔSR3	ΔTF4	ΔSR4	ΔTF5	ΔSR5
2022	-76%	-63%	-78%	-63%	-84%	-72%	-82%	-72%	-78%	-39%
2021	-32%	-24%	-56%	-54%	-62%	-56%	-60%	-53%	-74%	-68%
2020	-22%	-2%	-19%	60%	-32%	0%	-30%	-5%	-40%	-3%
2019	-22%	-13%	-39%	-25%	-21%	-10%	-25%	1%	-33%	-13%
2018	-5%	18%	-5%	26%	0%	26%	-6%	11%	-8%	7%
2017	-17%	-12%	-25%	-19%	-18%	-26%	-12%	-23%	-15%	-66%
2016	6%	0%	23%	30%	-7%	13%	-6%	12%	-1%	23%
2015	-13%	24%	-8%	28%	16%	24%	13%	12%	10%	27%
2014	-4%	-12%	-10%	-2%	-15%	-5%	-12%	3%	-12%	5%
2013	-13%	-19%	-10%	-15%	-2%	-7%	4%	0%	2%	0%
Avg	***-20%***	***-10%***	***-23%***	***-3%***	***-22%***	***-11%***	***-22%***	***-12%***	***-25%***	***-13%***

In Table 4.5 the predictions under the simple regression method appear to yield better results in average terms for all periods. However, when examined individually, except for 2016 and 2017, the tree-fit method provides better predictions. In these years, the prices exhibited volatility, fluctuating up and down. As a result, the pattern seemed to change, but the tree-fit method remained consistent in its approach and presumably leveraged its historical learning to produce better results.

In Table 4.6 the predictions under the tree-fit method appear to yield better results in average terms for all periods except Period 3 and Period 5. Individually speaking,

Table 4.6: Price Prediction Preciseness Between Models_Corn

Years	**ΔTFW1**	**ΔSR1**	**ΔTF2**	**ΔSR2**	**ΔTF3**	**ΔSR3**	**ΔTF4**	**ΔSR4**	**ΔTF5**	**ΔSR5**
2022	-80%	-62%	-80%	-44%	-81%	-45%	-77%	-41%	-70%	84%
2021	-37%	-45%	-52%	-52%	-60%	-60%	-60%	-50%	-78%	-74%
2020	-3%	-13%	-7%	16%	-6%	3%	-13%	43%	-35%	-37%
2019	-1%	-9%	-23%	-32%	-8%	-8%	-46%	-62%	-2%	-5%
2018	17%	16%	-2%	-9%	4%	10%	-14%	-30%	12%	28%
2017	9%	-50%	-6%	-65%	4%	-25%	10%	40%	9%	37%
2016	16%	221%	5%	84%	4%	86%	21%	156%	-2%	50%
2015	7%	9%	1%	119%	11%	46%	20%	47%	18%	13%
2014	-1%	2%	-18%	-59%	-12%	31%	-1%	24%	1%	2%
2013	2%	93%	34%	188%	-6%	89%	27%	209%	-5%	-6%
Avg	***-7%***	***16%***	***-15%***	***15%***	***-15%***	***13%***	***-13%***	***34%***	***-15%***	***9%***

tree fit seems to exhibit more precision once again. When considering the volatility of corn prices, it seems that the tree-fit method provides better accuracy when there is price volatility.

4.2 Conditions and Outcomes

4.2.1 Highlights of the Study

The study provides price predictions based on climate conditions, specifically total precipitation and dew point, using a random forest decision tree-fit modeling approach, which is essentially a sophisticated machine learning model. These predictions can then be integrated into decision tree arguments alongside other decision layers.

The price predictions demonstrate coherence when the tree-fit model is employed. Therefore, unless there is a significant inflation or exchange rate difference, the price predictions are expected to be accurate. However, it's important to consider other factors such as comprehensive information on the total planted area and more precise import-export figures, as incorporating these factors would likely improve the accuracy of the predictions.

4.2.2 Strength of the Study

This research stands out from other studies linking climate and price because it utilizes ECMWF data. Unlike manually measured and recorded weather data, the data used here consists of 2-meter precipitation and near-surface air temperature figures (dew point), which are calculated. This ensures that the data is precise and comparable since it originates from a standardized source, eliminating potential human errors.

Furthermore, the data is represented at specific points with a longitude and latitude grid of [0.1, 0.1], rather than encompassing a larger area with varying acceptable perimeters determined by an observer station. This results in more precise and dense data, enhancing the accuracy of the analysis.

By linking climate conditions to price level predictions, this study takes models that only link climate conditions and yield one step further. It expands the scope of analysis beyond just yield predictions, offering insights into how climate influences market prices, thereby providing valuable information for decision-making in agricultural and economic sectors.

4.2.3 Policy Implications

This study;

1. Utilizes weather inputs of precipitation and dew point, as these inputs are **primary** and **pivotal** weather conditions **as outlined in Chapter 2, "Required Weather Conditions for Production" sections inputs**, also they are accessible and feasible to be accurately measured and recorded as in the **nowcast data archive of ECMWF**,

2. For predicting future prices of wheat, barley and corn, as they are **majorly produced and traded grains** so that **the study output could be meaningful to cover the rest of the grains by their representation percentage**.

3. By using an **advanced machine learning technique known as the ensemble tree learning (referenced as "tree-fit" in this paper) method**, **as this method can efficiently be applied with two blocks of data among other machine learning methods mentioned in Section 3.2.1.1**.

These inputs and methods listed above shows that the study covers almost every conceivable aspect, spanning a widest scope and the largest time interval available so holds promise for paving the way for subsequent scholars to enhance price forecasting in other agricultural commodities.

On the output side;

1. This study is believed **to improve the confidence in trading process**.
    - Understanding **the future reference price generated by this model would help traders and financial market investors to avoid being misled by speculation and to discourage them to set lower sell rates and higher buy rates than the market norms**.
    - By undertaking this approach, unnecessary **volatility can be mitigated, thereby promoting stable economic growth**.
    - Asymmetric information would be reduced, leading to improved fairness in market conditions and **greater encouragement for farmers and merchants to participate in trading, thereby enhancing market sophistication at all levels**.
    - Diversification of financial instruments would accommodate investors at varying risk levels, thereby **boosting trading volume** as well as the output would have a potential to trigger further financial innovations in agricultual sector.
2. As for policy side, accurately predictiong future product prices **aids policymakers in discerning the necessity for public interventions and subsidies**. As mentioned in Section 2.1.5, TMO, serving as market maker, facilitates public subsidies for certain products having a big impact on public nutrition and also the grain market. By determining the prices on buy and sell side, this system enables efficient monitoring of subsidies and evaluation of agricultural policies [47] [25]. Among the grain types used in the study, wheat occupies a prominent role, with buying rates accounting for roughly 15% of overall output on average. For barley, this figure stands at 5%, and for corn, it is 12% [48]. TMO implements these strategies to stabilize the market and guide farmers' planting decisions, aiming for sustainable and efficient harvests in future

years. The outcome of this study is expected to serve **as a robust demonstrator for making decisions based on a reference price known in a year advance, allowing market participants to position themselves accordingly**. This approach would minimize disputes and noise, facilitating fair support and enabling physical trading to occur smoothly.

3. Similarly, merchants and entrepreneurs planning to produce processed food can utilize this future reference price output of this study **to allocate their resources for investment, enabling efficient production planning and minimizing idle resources**.

4. The study's output, specifically the future reference price, will serve as **a crucial indicator for insurance companies in developing their risk models** also encouraging risk averse stakeholders to participate in insurance market. These future reference prices will aid **in indemnity calculations**, allowing insurance premiums to be adjusted accordingly and ensuring that losses are fairly calculated to cover farmers' damages.

5. Last but not least, constructing a price index based on the future reference price from **this study's output would benefit the forthcoming future commodity market in Turkey**. This commodity market could use the index as an underlying asset for index-based derivative products, thereby **creating a new market instrument**.

4.2.4 Limitations of the Study

Several limitations were encountered in this study regarding scope, applicability, and data diversity. While the data was sourced from reliable sources, the study's scope may be perceived as narrow. However, efforts were made to compensate for this by enhancing representation effectiveness. These limitations can serve as valuable insights for future researchers, aiding in the improvement of results while upholding the validity and credibility of the initial study.

- Diversity of meteorological events: Despite aiming to link the final outcome with climate conditions, not all meteorological events were considered due to the lack of measurable and consistent data. Variables such as wind, floods, and meteorological shocks were not directly incorporated but were implicitly represented through dew point and total precipitation.

- Scope of land: The study focused on only five provinces instead of examining the entire planted area. To address this limitation, the selected provinces were chosen based on their representation power, focusing on regions where the products are most traded and produced, and which exhibit diverse climate conditions.

- The latitude and longitude of the area: As the provinces are not perfectly rectangular in shape, the boundaries were determined as wide as possible. However, some areas may still fall outside the defined boundaries. While this may not significantly impact the results as weather conditions are likely to be similar over minor distances, it remains a consideration.

- Data availability: The availability of monthly local prices was limited to data after 2010. This discrepancy affected the models, as the yield model encompassed 23 years while the price model could only consider 13 years. Additionally, some price data, such as those from the Polatlı Local Exchange, were recorded manually, which may introduce inconsistencies. However, efforts were made to ensure the models provided consistent results.

Furthermore, the yield data for Polatlı was not available, so the yield data for Ankara was used as a substitute. However, the climate conditions in Ankara were deemed representative of the broader region, as evidenced by the isophane maps in figures Figure 2.1, Figure 2.2 and Figure 2.3, and the periods were selected with Ankara's climate in mind.

Finally, there may be slight discrepancies in export and import figures due to variations in available custom codes for the products across different years. Efforts were made to align the numbers as closely as possible, and details regarding the selected custom codes and related information are provided in the appendix Section A.5.

4.3 Conclusion

In this study, the accurate prediction of prices for three primary grains is pursued through appropriate modeling techniques. To achieve continuous and highly accurate predictions, a random forest approach is employed, utilizing ensemble regression for machine learning. This process is executed in two phases:

- Phase 1: Weather forecasts, represented by total precipitation and dew point figures, serve as independent variables, while the yield of a specific grain in a particular province acts as the dependent variable.
- Phase 2: Yields in selected provinces are considered independent variables, along with additional factors such as net import and global price, to predict the local price of the respective grain accurately.

To assess the acceptability of the predictions, another method, namely simple regression prediction, is also applied. Results indicate that in Phase 1, where the environment is more complex, simple regression prediction performed poorly, while the tree-fit model exhibited a stable trajectory despite slightly lower accuracy compared to Phase 2. Conversely, in Phase 2, simple regression yielded better results overall, but the tree-fit model demonstrated greater accuracy in volatile environments. Notably, the tree-fit model produced reliable predictions even with lower statistical significances (see also Table 3.5).

In conclusion, among ensemble regression methods, the tree-fit training method proved effective for price predictions, particularly in scenarios requiring extensive data and layered analysis. While it may perform optimally in stable economic conditions rather than environments with high inflation or exchange rate volatility, this method holds promise for forecasting future prices in commodity markets, thereby facilitating more informed decision-making for producers and financial instrument traders. By leveraging this approach, it is possible to anticipate future price indices, allowing for the determination of preferred weights by index structurers. Ultimately, this thesis study serves as a valuable reference for researchers seeking to empower producers with efficient production choices and enable financial traders to establish reasonable pricing, enlighten insurance companies for their risk mappings, thus mitigating volatility in commodity markets and fostering sustainable agricultural supply chains.

REFERENCES

[1] T. Akar, M. Avci, and D. F., Barley: Post-harvest operations, 2004, https://openknowledge.fao.org/handle/20.500.14283/au997e, Access date: 29.05.2024.

[2] I. Akkaya and E. Hafizoglu, *Buğday Yetiştiriciliği ve Yeni Buğday Çeşitleri*, Bursa Valiliği, İl Tarım Müdürlüğü, Çiftçi Eğitimi ve Yayım Şube Müdürlüğü, ÇEY 2009/II, pp. 4-14, 2009, https://kutuphane.tarimorman.gov.tr/vufind/Record/19015, Access date: 27.05.2024.

[3] M. Babaoglu, *Mısır Tarımı*, Trakya Tarımsal Araştırma Enstitüsü Müdürlüğü, 2024, https://arastirma.tarimorman.gov.tr/ttae/Sayfalar/Detay.aspx?SayfaId=89, Access date: 29.05.2024.

[4] Bitki Sağlığı ve Karantina Daire Başkanlığı, *Hububat Hastalık ve Zararlıları ile Mücadele*, T.C. Gıda, Tarım ve Hayvancılık Bakanlığı Gıda ve Kontrol Genel Müdürlüğü, pp. 9-15, 2016, https://www.tarimorman.gov.tr/GKGM/Belgeler/Uretici_Bilgi_Kosesi/Dokumanlar/hububat.pdf, Access date: 27.05.2024.

[5] L. Breiman, Random forests, Machine Learning, 45, pp. 5–32, 2001.

[6] R. Cengiz, *Mısır Yetiştiriciliği*, T.C. Gida Tarim ve Hayvancilik Bakanligi, ÇEY 2012/9, 2009, https://kutuphane.tarimorman.gov.tr/pdf_goster?file=92de9eaf0fd6af0ce2ccacc9e7abd8b3#BOOK/, Access date: 27.05.2024.

[7] Central Bank of Republic of Türkiye, USD/TL exchange rate, https://www.tcmb.gov.tr/kurlar/kurlar_tr.html, Access date: 27.03.2024.

[8] CME, CME Corn Contract Specs, https://www.cmegroup.com/markets/agriculture/grains/corn.contractSpecs.html, Access date: 27.05.2024.

[9] CME, CME Wheat Contract Specs, https://www.cmegroup.com/markets/agriculture/grains/wheat.contractSpecs.html, Access date: 27.05.2024.

[10] Coğrafi Bilgi Sistemleri ve Uzaktan Algılama Bölümü, *Agro-Ekolojik Alt Bölgeler Atlası*, Tarla Bitkileri Merkez Araştırma Enstitüsü, 2009.

[11] T. G. Dietterich, Ensemble Methods in Machine Learning, 2000, https://doi.org/10.1007/3-540-45014-9_1 Access date: 27.05.2024.

[12] ECMWF, Climate Data Store, Total Precipitation and Dew Point Data, https://cds.climate.copernicus.eu/cdsapp#!/toolbox, Access date: 27.03.2023.

[13] S. Egilmez, *Ürün Raporu Arpa*, TEPGE, 2022, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20%C3%9Cr%C3%BCn%20Raporlar%C4%B1/2022%20%C3%9Cr%C3%BCn%20Raporlar%C4%B1/Arpa%20%C3%9Cr%C3%BCn%20Raporu%20TEPGE-2022%20361.pdf, Access date:29.05.2024.

[14] S. Egilmez, *Ürün Raporu Arpa 2023*, TEPGE AEPDI, 2023, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20%C3%9Cr%C3%BCn%20Raporlar%C4%B1/2023%20%C3%9Cr%C3%BCn%20Raporlar%C4%B1/Arpa%20%C3%9Cr%C3%BCn%20Raporu%202023-386%20TEPGE.pdf, Access date: 30.05.2024.

[15] FAO, Crop information/ maize, https://www.fao.org/land-water/databases-and-software/crop-information/maize/en/, Access date: 29.05.2024.

[16] FAO, Crop information/ wheat, https://www.fao.org/land-water/databases-and-software/crop-information/wheat/en/, Access date: 29.05.2024.

[17] Federal Reserve Bank of St Louis, FRED Graph Observations, India Consumer Price Index, https://fred.stlouisfed.org/series/INDCPIALLMINMEI, Access date: 27.03.2024.

[18] Federal Reserve Bank of St Louis, FRED Graph Observations, USA Consumer Price Index, https://fred.stlouisfed.org/series/CUUR0000SAF11, Access date: 27.03.2024.

[19] International Grains Council, Goi index, https://www.igc.int/en/markets/marketinfo-goi.aspx, Access date: 04.05.2024.

[20] Investing.com, Barley Futures,NBLc1, https://in.investing.com/commodities/ncdex-barley-futures-historical-data, Access date: 04.06.2024.

[21] Y. Jiao, H. Chen, H. Han, and Y. Chang, Development and utilization of corn processing by-products: A review, Foods, 11(22), pp. 1–15, 2022, https://doi.org/10.3390/foods11223709, Access date:29.05.2024.

[22] N. Kara, N. Yıldız, and S. Keskin, *Buğday - Arpa Tarımı*, T.C. Gıda Tarım ve Hayvancılık Bakanlığı, ÇEY 2012/3, pp. 7-12, 2012, https://

kutuphane.tarimorman.gov.tr/vufind/Record/1178774, Access date: 27.05.2024.

[23] Y. Kaya and M. Akcura, İklim değişiminin farklı büyüme tabiatına sahip bazı arpa Çeşitleri Üzerine etkisi, Türk Tarım ve Doğa Bilimleri Dergisi, 9(1), pp. 107–118, 2022, https://doi.org/10.30910/turkjans.1017266 Access date:29.05.2024.

[24] B. Kocarik Gacar and I. Deveci Kocakoc, Regression analyses or decision trees?, Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi, 18(4), pp. 251–260, 2020.

[25] Kredi Kayıt Bürosu, Çiftçi kayıt sistemi, https://www.kkb.com.tr/urunler/ciftci-kayit-sistemi, Access date: 30.05.2024.

[26] MATLAB, What Is a Machine Learning Model?, Types of Machine Learning Models, https://www.mathworks.com/discovery/machine-learning-models.html, Access date: 30.05.2024.

[27] S. Mittal, *Wheat and Barley Production Trends and Research Priorities: A Global Perspective*, Springer Nature Singapore Pte Ltd., pp. 3-16, 2022, https://doi.org/10.1007/978-981-16-4449-8, Access date: 27.05.2024.

[28] Morgan Stanley Capital International, MSCI Index Calculation Methodologies, https://www.msci.com/eqb/methodology/meth_docs/MSCI_Mar2011_IndexCalcMethodology.pdf, Access date:30.05.2024.

[29] NCDEX, NCDEX Barley Contract Specs, https://www.ncdex.com/products/BARLEYJPR, Access date:02.05.2024.

[30] A. Ozcelik and O. O. Ozer, Koyck Modeliyle Türkiye'de Buğday Üretimi ve Fiyatı İlişkisinin Analizi, Journal of Agricultural Sciences, 12(04), pp. 333–339, 2022.

[31] A. Ozdemir, A. Onan, and V. Cinarli Ergene, Machine Learning and Ensemble Learning Based Method Using Online Employee Assessments to Identify and Analyze Job Satisfaction Factors, European Journal of Science and Technology, 40, pp. 19–28, 2022.

[32] K. Polat, *Buğday Durum Raporu, 2022 July*, TEPGE AEPDI, 2022, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasalar%C4%B1/2022-Temmuz%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Raporu/7-BU%C4%9EDAY%20T%C3%9CP%20Temmuz%202022.pdf, Access date: 27.05.2024.

[33] K. Polat, *Durum Tahmin BUĞDAY 2023*, TEPGE AEPDI, 2023, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF\%20Tar\%C4\%B1m\%20\%C3\%9Cr\%C3\%BCnleri\%20Piyasalar\%C4\%B1/2022-Temmuz\%20Tar\%C4\%B1m\%20\%C3\%9Cr\%C3\%BCnleri\%20Raporu/MISIR\%20T\%C3\%9CP\%20TEMMUZ\%202022.pdf, Access date: 30.05.2024.

[34] Polatlı Ticaret Borsası, Günlük tescil büteni, https://www.polatliborsa.org.tr/gunluk-tescil-bulteni/, Access date: 04.06.2024.

[35] Reserve Bank of India, USD/INR exchange rate, https://www.rbi.org.in/scripts/referenceratearchive.aspx, Access date: 27.03.2024.

[36] M. Reynolds and H. Braun, *Wheat Improvement, Food Security in a Changing Climate*, Springer Cham, 2022, https://doi.org/10.1007/978-3-030-90673-3, Access date: 27.05.2024.

[37] J. Sexton and P. Laake, Standard errors for bagged and random forest estimators, Computational Statistics and Data Analysis, 53(3), pp. 801–811, 2008.

[38] O. Simsek, G. Yucel, Y. Nadaroglu, O. Dokuyucu, and S. Gokdag, *Türkiye Fenoloj Atlası*, Meteoroloji Genel Müdürlüğü, Araştırma Dairesi Başkanlığı, Zirai Meteoroloji Şube Müdürlüğü, 2014, https://cevresehiriklimkutuphanesi.csb.gov.tr/BOOKs/BOOK/53d6c7d7-be22-4bc5-961c-c36c4910da87, Access date: 27.05.2024.

[39] S&P Global Commodity Insights, *Specification Guide, Global Grains and Oilseeds*, Platts, 2022, https://www.spglobal.com/commodityinsights/PlattsContent/_assets/_files/en/our-methodology/methodology-specifications/global-grains-oilseeds.pdf, Access date: 23.02.2023.

[40] TARSİM, Crop insurance, https://www.tarsim.gov.tr/eng/insurances/insurance-crop.jsp, Access date: 17.04.2023.

[41] K. Tasdan, *Mısır Durum Raporu, 2022 Temmuz*, TEPGE AEPDI, 2022, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasalar%C4%B1/2022-Temmuz%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Raporu/7-BU%C4%9EDAY%20T%C3%9CP%20Temmuz%202022.pdf, Access date: 27.05.2024.

[42] K. Tasdan, *Mısır Durum Raporu, 2023 Haziran*, TEPGE AEPDI, 2023, https://arastirma.tarimorman.gov.tr/tepge/

Belgeler/PDF%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasalar%C4%B1/2023-Temmuz%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Raporu/MISIR%20T%C3%9CP%20HAZ%C4%B0RAN%202023-TEPGE.pdf, Access date: 30.05.2024.

[43] TEPGE AEPDI, Durum tahmin misir 2023, 2003, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF\%20Durum-Tahmin\%20Raporlar\%C4\%B1/2023\%20Durum-Tahmin\%20Raporlar\%C4\%B1/M\%C4\%B1s\%C4\%B1r\%20Durum\%20Tahmin\%20Raporu\%202023-390\%20TEPGE.pdf, Access date: 30.05.2024.

[44] TMEX, Peridoical Bulletin, https://www.turib.com.tr/en/periodical-bulletin/, Access date: 04.06.2024.

[45] TMEX, Storage Periods of Agricultural Products in Turkey, https://www.turib.com.tr/en/period-of-storage-for-commodities/, Access date: 27.05.2024.

[46] TMEX, TMEX Agricultural Commodity Indices, https://www.turib.com.tr/en/indices/, Access date: 27.05.2024.

[47] TMO, Turkish Grain Board, https://www.tmo.gov.tr/homepage, Access date: 30.05.2024.

[48] TMO, Turkish Grain Board Annual Purchase Rate Over Production, https://www.tmo.gov.tr/bilgi-merkezi/tablolar, Access date: 02.06.2024.

[49] TURKSTAT, Turkish Agricultural Production, https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr, Access date: 27.03.2024.

[50] TURKSTAT, Turkish Annual Foreign Trading, https://biruni.tuik.gov.tr/disticaretapp/menu.zul, Access date: 27.03.2024.

[51] TURKSTAT, Turkish Consumer Price Index, https://biruni.tuik.gov.tr/medas/?kn=159&locale=tr, Access date: 27.03.2024.

[52] TURKSTAT News letter, Devlet Destekli Tarım Sigortaları, 2021, https://data.tuik.gov.tr/Bulten/, Access date: 27.05.2024.

[53] US Grains Council, Converting grain units, https://grains.org/markets-tools-data/tools/converting-grain-units/, Access date: 27.05.2024.

[54] G. Verma, M. Barman, and S. K. Singh, *Wheat: Production, Protection and Improvement*, Satish Serial Publishing House, 2022.

[55] Yahoo Finance, CBOT - Delayed Quote, USX, Chicago SRW Wheat Futures, https://finance.yahoo.com/quote/ZW%3DF/history/?period1=963878400&period2=977097600&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true, Access date: 04.06.2024.

[56] Yahoo Finance, CBOT- Delayed Quote, USX, Corn Futures, https://finance.yahoo.com/quote/ZC%3DF/history/?period1=963878400&period2=964137600&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true, Access date: 04.06.2024.

APPENDIX A

MODEL INPUTS

A.1 Data Set and Intervals

The data set periods and sources are shown in Table A.1 as mentioned in Section 3.1.2.

Table A.1: Provided Data and Their Sources

Domestic Prices	**Period**	**Source**
Wheat	2010-2019	Konya Mercantile Exchange – via email
Wheat	2010-2013; 2015-2019	Yozgat Mercantile Exchange – via email
Wheat	2019	Polatlı Mercantile Exchange – daily bulletin [34]
Wheat	2020, 2011, 2022	TMEX – Periodical Bulletin [44]
Barley	2020, 2011, 2022	Konya Mercantile Exchange – via email
Barley	2010-2013; 2015-2019	Yozgat Mercantile Exchange – via email
Barley	2019	Polatlı Mercantile Exchange – daily bulletin [34]
Barley	2020, 2011, 2022	TMEX – Periodical Bulletin [44]
Corn	2010-2019	Konya Merchantile Exchange – via email
Corn	2017-2022	Sanliurfa Merchantile Exchange – via email
Corn	2020, 2011, 2022	TMEX – Periodical Bulletin [44]
Global Prices	**Period**	**Source**
Wheat Futures	2010-2022	CBOT– Yahoo Finance [55]
Barley Futures	2010-2022	NCDEX– Yahoo Finance [20]
Corn Futures	2010-2022	CBOT – Yahoo Finance [56]
Inflation Rate	**Period**	**Source**
Turkey	2010-2022	TURKSTAT [51]
India	2010-2022	FED Economic Data [17]
USA	2010-2022	FED Economic Data[18]

Table A.1: Provided Data and Their Sources (continued)

Exchange rates	**Period**	**Source**
USD/TRY	2010-2022	CBT [7]
INR/USD	2010-2022	Reserve Bank of India[35]
Import and Export	**Period**	**Source**
Wheat	2010-2022	TURKSTAT (Custom Codes:100199,100119)[50]
Barley	2010-2022	TURKSTAT (Custom Code:100390) [50]
Corn	2010-2022	TURKSTAT (Custom Code:100590) [50]
Domestic Production	**Period**	**Source**
Wheat, Barley and Corn Yield in Konya, Ankara, Yozgat, Adana and Sanliurfa	2000-2022 [1]	TURKSTAT [49]
Total Precipitation	**Period**	**Source**
Konya, Polatlı, Yozgat, Adana and Sanliurfa	2000-2022	ECMWF[12]
Dew Point	**Period**	**Source**
Konya, Polatli, Yozgat, Adana and Sanliurfa	2000-2022	ECMWF[12]

A.2 Depedent Variable Phase 1: Annual Yield Based on Provinces

Yield data is available province base for years between 2004-2022 in TURKSTAT database [49]. For years 2000-2003, only Turkey's total yield is available. Therefore, "the province yield/total yield" ratio per year among 2004-2022 is taken and the average of those ratios is applied individually to annual total yield to get province based yield figure (see also Table A.1). Independent varibales are taken from ECMWF [12] and the scores are derived according to the steps explained in Section 3.2.2.1 so they are not included in this Appendix section.

Below, Table A.2 for wheat, Table A.3 for barley and Table A.4 for corn shows the studied figures, accordingly.

[1] The province-based production figures are available for the period 2004-2022. For the years 2000-2003, province-based productions were estimated by calculating the average production ratio for each province and applying it to the total production figure for the respective year published by TURKSTAT.

Table A.2: Annual Wheat Yield Per Province in Model Phase 1

Year/ Kg	**W Konya**	**W Ankara**	**W Yozgat**	**W Adana**	**W Şanlıurfa**
2022	1,929,537	896,418	669,114	544,186	1,007,891
2021	1,579,839	609,592	437,245	707,521	1,182,655
2020	1,920,700	974,528	606,758	621,849	1,538,904
2019	1,886,131	1,053,032	553,601	502,562	677,390
2018	2,037,936	1,093,264	670,513	681,905	860,601
2017	2,192,410	1,090,500	699,052	690,411	1,044,645
2016	2,045,298	1,205,676	734,593	621,872	917,545
2015	2,554,256	1,150,555	830,939	730,873	1,087,746
2014	1,905,300	817,110	644,926	583,417	867,558
2013	2,291,930	1,153,980	776,657	704,481	1,215,004
2012	1,570,660	1,026,700	759,737	785,352	1,121,393
2011	2,444,814	1,141,228	927,056	759,020	1,037,447
2010	1,515,303	936,340	800,283	709,017	974,612
2009	1,696,165	1,156,461	821,929	788,778	1,189,346
2008	1,089,782	783,221	613,056	899,461	615,189
2007	1,026,565	539,788	588,114	903,584	850,183
2006	1,586,033	1,137,404	658,839	855,523	925,294
2005	1,343,977	1,208,484	818,838	1,014,820	996,585
2004	1,497,569	1,035,230	659,731	836,922	874,923
2003	1,684,153	937,924	656,948	696,855	942,650
2002	1,728,473	962,606	674,236	715,193	967,457
2001	1,684,153	937,924	656,948	696,855	942,650
2000	1,861,432	1,036,653	726,101	770,208	1,041,877

Table A.3: Annual Barley Yield Per Province in Model Phase 1

Year/ kg	**B Konya**	**B Ankara**	**B Yozgat**	textbfB Adana	**B Şanlıurfa**
2022	1,264,821	725,022	158,294	28,911	167,302
2021	843,102	408,098	108,417	26,082	126,819
2020	1,266,362	812,919	216,085	24,226	226,295
2019	1,146,786	649,091	103,813	14,289	387,507
2018	968,554	572,080	87,062	14,796	339,289
2017	873,016	567,898	84,341	12,025	530,669
2016	809,258	665,165	85,120	12,118	243,298
2015	972,570	688,728	111,34	14,243	588,764
2014	732,799	459,764	95,150	12,003	334,670
2013	838,491	683,764	97,552	11,725	613,385
2012	706,837	550,969	93,372	13,665	581,161
2011	929,071	687,039	108,696	14,590	448,675
2010	653,978	706,081	124,863	13,971	485,851
2009	838,496	683,264	179,282	12,894	582,246
2008	515,501	413,031	158,392	13,676	300,694
2007	606,630	285,099	193,354	16,264	745,173
2006	1,126,169	740,198	247,959	20,032	711,695
2005	876,827	779,510	250,241	30,222	754,106
2004	939,957	659,014	223,707	25,296	578,414
2003	953,499	659,403	151,410	18,623	487,046
2002	977,043	675,685	155,149	19,083	499,072
2001	882,870	610,559	140,195	17,244	450,969
2000	941,728	651,263	149,541	18,393	481,033

Table A.4: Annual Corn Yield Per Province in Model Phase 1

Year/ kg	C Konya	C Ankara	C Yozgat	C Adana	C Şanlıurfa
2022	2,044,202	62,816	55	888,348	870,822
2021	1,261,475	43,610	3,376	810,145	819,764
2020	1,070,626	32,815	860	819,978	1,049,849
2019	1,345,064	33,397	248	717,802	354,710
2018	1,104,538	11,536	58	842,697	219,384
2017	621,884	7,444	62	1,036,130	422,950
2016	638,3	6,559	82	1,086,606	547,715
2015	558,19	6,453	111	1,015,428	687,598
2014	382,099	7,381	161	1,005,651	581,560
2013	353,552	3,877	208	915,284	732,125
2012	312,059	2,217	37	682,462	608,991
2011	159,858	3,198	37	760,744	375,414
2010	103,43	2,925	34	748,160	453,006
2009	104,129	3,268	105	894,099	328,582
2008	80,307	3,652	128	1,036,626	186,078
2007	57,902	2,792	140	1,013,099	115,564
2006	51,577	3,324	175	1,014,235	88,013
2005	78,199	3,763	315	1,014,668	101,702
2004	76,547	2,247	207	722,394	49,536
2003	246,663	5,679	162	506,575	223,086
2002	184,997	4,259	122	379,931	167,314
2001	193,806	4,462	128	398,023	175,282
2000	202,616	4,665	133	416,115	183,249

A.3 Dependent Variable Phase 2: Local Real Prices

After raw data of local prices are taken from the sources stated in Table A.1, the below steps are taken to derive Local Real Prices.

- The prices taken from sources are grouped in monthly averages,
- The average of the available price of different sources for the same month is taken,
- The CPI Turkey [51] is applied and current prices are converted into real prices by keeping 2010 January prices same,
- The price of the month at the period end is taken to represent related period's price.
- Period ends are respectively: March (P 1), May (P 2), Jul (P 3), Sep (P 4) and Dec (P 5).

- At below Table A.5 for wheat, Table A.6 for barley and Table A.7 for corn shows the studied figures, accordingly.

Table A.5: Local Wheat Real Prices used in Model Phase 2

Year/TL	**Local Wh_P 1**	**Local Wh_P 2**	**Local Wh_P 3**	**Local Wh_P 4**	**Local Wh_P 5**
2022	3.2247	3.6655	4.4613	3.9963	3.6670
2021	1.0532	1.1644	1.2603	1.4549	2.8231
2020	0.7778	0.8157	0.8422	0.9512	1.0992
2019	0.6826	0.7775	0.6464	0.6806	0.8054
2018	0.5466	0.5159	0.4691	0.5301	0.5741
2017	0.5910	0.6267	0.5609	0.5362	0.5592
2016	0.5836	0.5918	0.5447	0.5729	0.5950
2015	0.6533	0.6238	0.5045	0.5351	0.5855
2014	0.6358	0.6252	0.6780	0.6492	0.6608
2013	0.5940	0.6039	0.5245	0.5199	0.5719
2012	0.4690	0.4952	0.5294	0.5584	0.5960
2011	0.5954	0.5718	0.5331	0.5882	0.4817
2010	0.4677	0.4601	0.4977	0.5324	0.6387

Table A.6: Local Barley Real Prices used in Model Phase 2

Year / TL	**Local Bar_P 1**	**Local Bar_P 2**	**Local Bar_P 3**	**Local Bar_P 4**	**Local Bar_P 5**
2022	3.0200	4.3132	4.3822	4.1888	3.8896
2021	0.9503	1.5172	1.6286	1.6046	2.7756
2020	0.7161	0.7041	0.7511	0.7571	0.9373
2019	0.7012	0.8865	0.6304	0.6797	0.8004
2018	0.5684	0.5640	0.4942	0.5381	0.5775
2017	0.6334	0.6877	0.5872	0.5677	0.6084
2016	0.4998	0.4319	0.5110	0.5266	0.5233
2015	0.5940	0.5699	0.4227	0.4458	0.4795
2014	0.5312	0.5681	0.5543	0.5551	0.5841
2013	0.5607	0.5476	0.4782	0.4760	0.5045
2012	0.5095	0.5221	0.5019	0.5401	0.5734
2011	0.4703	0.4637	0.4412	0.4560	0.4621
2010	0.3322	0.3412	0.3707	0.4171	0.4874

Table A.7: Local Corn Real Prices used in Model Phase 2

Year/TL	**Local Cor_P 1**	**Local Cor_P 2**	**Local Cor_P 3**	**Local Cor_P 4**	**Local Cor_P 5**
2022	2.8541	3.7127	3.9699	3.5036	2.7415
2021	0.7630	1.2124	1.4240	1.5257	2.2084
2020	0.5164	0.5778	0.6026	0.6129	0.6812
2019	0.5056	0.6795	0.6125	0.9158	0.4489
2018	0.4399	0.5331	0.5430	0.5646	0.3983
2017	0.4772	0.5516	0.5468	0.4485	0.4136
2016	0.4592	0.5001	0.5481	0.4210	0.4564
2015	0.5043	0.5185	0.5253	0.4419	0.3936
2014	0.5441	0.6066	0.6427	0.5329	0.4607
2013	0.5306	0.4101	0.5912	0.4505	0.4805
2012	0.4962	0.4906	0.5012	0.4937	0.4701
2011	0.5874	0.6165	0.6106	0.6639	0.4452
2010	0.4430	0.4322	0.4709	0.4776	0.4891

A.4 Phase 2 Independent Variable #1: Total Yield of Selected 5 Provinces

The total yield in Konya, Ankara, Yozgat, Adana and Urfa is taken from TURKSTAT as stated in Table A.1 and after the process mentioned in Section A.2 annual yield is divided into 12 to get monthly average and they are accumulated according to the number of month in that period to get periodical accumulated yield figures.

Periods are respectively: Jan-March (P 1), Apr-May (P 2), Jun-Jul (P 3), Aug-Sep (P 4) and Oct-Dec (P 5).

At below Table A.8 for wheat, Table A.9 for barley and Table A.10 for corn shows the studied figures, accordingly.

Table A.8: Wheat Yield Figures used in Model Phase 2

Year/ kg	**Y Wh_P 1**	**Y Wh_P 2**	**Y Wh_P 3**	**Y Wh_P 4**	**Y Wh_P 5**
2022	1,261,787	841,191	841,191	841,191	1,261,787
2021	1,129,213	752,809	752,809	752,809	1,129,213
2020	1,415,685	943,790	943,790	943,790	1,415,685
2019	1,168,179	778,786	778,786	778,786	1,168,179
2018	1,336,055	890,703	890,703	890,703	1,336,055
2017	1,429,255	952,836	952,836	952,836	1,429,255
2016	1,381,246	920,831	920,831	920,831	1,381,246
2015	1,588,592	1,059,062	1,059,062	1,059,062	1,588,592
2014	1,204,578	803,052	803,052	803,052	1,204,578
2013	1,535,513	1,023,675	1,023,675	1,023,675	1,535,513
2012	1,315,961	877,307	877,307	877,307	1,315,961
2011	1,577,391	1,051,594	1,051,594	1,051,594	1,577,391
2010	1,233,889	822,593	822,593	822,593	1,233,889

Table A.9: Barley Yield Figures used in Model Phase 2

Year/ kg	**Y Bar_P 1**	**Y Bar_P 2**	**Y Bar_P 3**	**Y Bar_P 4**	**Y Bar_P 5**
2022	586,088	390,725	390,725	390,725	586,088
2021	378,130	252,086	252,086	252,086	378,130
2020	636,472	424,315	424,315	424,315	636,472
2019	575,372	383,581	383,581	383,581	575,372
2018	495,445	330,297	330,297	330,297	495,445
2017	516,987	344,658	344,658	344,658	516,987
2016	453,740	302,493	302,493	302,493	453,740
2015	593,911	395,941	395,941	395,941	593,911
2014	408,597	272,398	272,398	272,398	408,597
2013	561,229	374,153	374,153	374,153	561,229
2012	486,501	324,334	324,334	324,334	486,501
2011	547,018	364,679	364,679	364,679	547,018
2010	496,186	330,791	330,791	330,791	496,186

Table A.10: Corn Yield Figures used in Model Phase 2

Year/ kg	Y Cor_P 1	Y Cor_P 2	Y Cor_P 3	Y Cor_P 4	Y Cor_P 5
2022	966,561	644,374	644,374	644,374	966,561
2021	734,593	489,728	489,728	489,728	734,593
2020	743,532	495,688	495,688	495,688	743,532
2019	612,805	408,537	408,537	408,537	612,805
2018	544,553	363,036	363,036	363,036	544,553
2017	522,118	348,078	348,078	348,078	522,118
2016	569,816	379,877	379,877	379,877	569,816
2015	566,945	377,963	377,963	377,963	566,945
2014	494,213	329,475	329,475	329,475	494,213
2013	501,262	334,174	334,174	334,174	501,262
2012	401,442	267,628	267,628	267,628	401,442
2011	324,813	216,542	216,542	216,542	324,813
2010	326,889	217,926	217,926	217,926	326,889

A.5 Phase 2 Independent Variable #2: Net Import (M-X) Figures of Turkey

The annual export and import figures are taken from TURKSTAT as stated in Table A.1 as per the below custom codes, export figures are deducted from import figures to get net import figures (M-X) and monthly net imports are accumulated periodically.

Periods are respectively: Jan-March (P 1), Apr-May (P 2), Jun-Jul (P 3), Aug-Sep (P 4) and Oct-Dec (P 5).

The available data in TURKSTAT for edible grains (since the prices are belonging to them) between the selected time interval was,

- Wheat [50]: "100199 - Wheat (except for durum wheat for pasta) and mixed; except for the ones for seed" and "100119 - Wheat for pasta (durum wheat); except for the ones for seed" between 2012-2020 and "100190 - Wheat and mixed (other)" for 2010 and 2011.
- Barley[50]: "100390 - Barley; except for the ones for seed" between 2012-2020 and "10030090 - Barley (other)" for 2010 and 2011.
- Corn [50]: "100590 - Corn; except for the one for seed" between 2010-2022

At below Table A.11 for wheat, Table A.12 for barley and Table A.13 for corn shows the studied figures, accordingly.

Table A.11: Wheat Net Import Figures used in Model Phase 2

Year/ kg	**MX Wh_P 1**	**MX Wh_P 2**	**MX Wh_P 3**	**MX Wh_P 4**	**MX Wh_P 5**
2022	2,204,166,548	1,469,444,366	1,469,444,366	1,469,444,366	2,204,166,548
2021	2,025,505,362	1,350,336,908	1,350,336,908	1,350,336,908	2,025,505,362
2020	2,429,934,661	1,619,956,441	1,619,956,441	1,619,956,441	2,429,934,661
2019	2,442,679,502	1,628,453,002	1,628,453,002	1,628,453,002	2,442,679,502
2018	1,437,539,170	958,359,447	958,359,447	958,359,447	1,437,539,170
2017	1,239,878,829	826,585,886	826,585,886	826,585,886	1,239,878,829
2016	1,055,374,363	703,582,909	703,582,909	703,582,909	1,055,374,363
2015	1,074,401,392	716,267,594	716,267,594	716,267,594	1,074,401,392
2014	1,312,118,667	874,745,778	874,745,778	874,745,778	1,312,118,667
2013	50,857,102	633,904,735	633,904,735	633,904,735	950,857,102
2012	904,011,400	602,674,267	602,674,267	602,674,267	904,011,400
2011	1,187,362,294	791,574,862	791,574,862	791,574,862	1,187,362,294
2010	345,796,586	230,531,057	230,531,057	230,531,057	345,796,586

Table A.12: Barley Net Import Figures used in Model Phase 2

Year/ kg	**MX Bar_P 1**	**MX Bar_P 2**	**MX Bar_P 3**	**MX Bar_P 4**	**MX Bar_P 5**
2022	357,552,069	238,368,046	238,368,046	238,368,046	357,552,069
2021	530,634,215	353,756,144	353,756,144	353,756,144	530,634,215
2020	221,695,513	147,797,009	147,797,009	147,797,009	221,695,513
2019	125,460,852	83,640,568	83,640,568	83,640,568	125,460,852
2018	162,861,485	108,574,323	108,574,323	108,574,323	162,861,485
2017	93,842,434	62,561,623	62,561,623	62,561,623	93,842,434
2016	9,010,519	6,007,012	6,007,012	6,007,012	9,010,519
2015	49,871,058	33,247,372	33,247,372	33,247,372	49,871,058
2014	166,686,953	111,124,635	111,124,635	111,124,635	166,686,953
2013	64,157,595	42,771,730	42,771,730	42,771,730	64,157,595
2012	-6,124,046	-4,082,698	-4,082,698	-4,082,698	-6,124,046
2011	8,281,934	5,521,289	5,521,289	5,521,289	8,281,934
2010	-110,758,051	-73,838,701	-73,838,701	-73,838,701	-110,758,051

Table A.13: Corn Net Import Figures used in Model Phase 2

Year/ kg	**MX Cor_P 1**	**MX Cor_P 2**	**MX Cor_P 3**	**MX Cor_P 4**	**MX Cor_P 5**
2022	734,383,190	489,588,793	489,588,793	489,588,793	734,383,190
2021	510,596,340	340,397,560	340,397,560	340,397,560	510,596,340
2020	514,233,018	342,822,012	342,822,012	342,822,012	514,233,018
2019	893,195,584	595,463,723	595,463,723	595,463,723	893,195,584
2018	517,866,077	345,244,051	345,244,051	345,244,051	517,866,077
2017	485,877,010	323,918,007	323,918,007	323,918,007	485,877,010
2016	125,704,657	83,803,105	83,803,105	83,803,105	125,704,657
2015	355,532,859	237,021,906	237,021,906	237,021,906	355,532,859
2014	342,735,450	228,490,300	228,490,300	228,490,300	342,735,450
2013	337,216,615	224,811,076	224,811,076	224,811,076	337,216,615
2012	198,516,264	132,344,176	132,344,176	132,344,176	198,516,264
2011	93,584,573	62,389,715	62,389,715	62,389,715	93,584,573
2010	111,858,565	74,572,376	74,572,376	74,572,376	111,858,565

A.6 Phase 2 Independent Variable #3: Global Prices

After raw data of global prices are taken from the sources stated in Table A.1, the below steps are taken to derive Global Real Prices and the period end prices are taken into consideration.

Period ends are respectively: March (1), May (2), Jul (3), Sep (4) and Dec (5).

- The prices taken from sources are grouped in monthly averages,
- For Wheat,
  - The contract prices are in bushel terms which is equal to 136 metric tons and in terms of US cents [53].
  - The prices are in divided into 136,000 to get kg terms in US cent and multiplied by 100 to get USD price.
  - They are converted into TL by taking CBT exchange rate [7] of the corresponding month average.
  - The US inflation rate [18] is applied to base month price and the result is deducted from current price and real price data is reached.
- For Barley,
  - The contract prices are in 10 metric tons per 100 rupee, making 100 kg per rupee.
  - The contract prices are divided into 100 to get kg terms per rupee.
  - Then they are converted to USD first [35], then USD prices are converted to Turkish Lira based on CBT exchange rate of the corresponding month average.
  - The Indian inflation rate [17]is applied to base month price and the result is deducted from current price and real price data is reached.
- For Corn,
  - The prices are in bushel terms which is equal to 127 metric tons and in terms of US cents [53].

- The prices are in divided into 127,000 to get kg terms in US cents and multiplied by 100 to get USD price.
- They are converted into TL by taking CBT exchange rate of the corresponding month average.
- The US inflation rate [18] is applied to base month price and the result is deducted from current price and real price data is reached.

• The price of the month at the period end is taken to represent related period's price.

At below Table A.14 for wheat, Table A.15 for barley and Table A.16 for corn shows the studied figures, accordingly.

Table A.14: Wheat Global Real Price Figures used in Model Phase 2

Year/TL	**Global Wh_P 1**	**Global Wh_P 2**	**Global Wh_P 3**	**Global Wh_P 4**	**Global Wh_P 5**
2022	11.9020	12.9455	10.1284	11.3104	10.0521
2021	3.4732	4.2577	4.1053	4.2885	7.7113
2020	2.4111	2.5372	2.5446	2.9396	3.3183
2019	1.7459	1.9695	2.0472	1.9470	2.2634
2018	1.2947	1.6175	1.7103	2.2979	1.9586
2017	1.0938	1.0664	1.2601	1.0541	1.1035
2016	0.9224	0.9398	0.8527	0.7929	0.9661
2015	0.8992	0.8878	1.0158	1.0013	0.9512
2014	1.0527	0.9876	0.7760	0.7432	0.9652
2013	0.8976	0.8908	0.8919	0.9136	0.8959
2012	0.8114	0.8024	1.1063	1.1197	1.0151
2011	0.8314	0.8595	0.7834	0.8586	0.7940
2010	0.5409	0.5396	0.6412	0.7731	0.8422

Table A.15: Barley Global Real Price Figures used in Model Phase 2

Year/TL	**Global Bar_P 1**	**Global Bar_P 2**	**Global Bar_P 3**	**Global Bar_P 4**	**Global Bar_P 5**
2022	3.8556	6.0846	6.4985	6.4990	6.7144
2021	1.3507	1.8930	1.8624	2.1889	3.9338
2020	1.1003	1.1027	0.9730	1.0803	1.1718
2019	1.0605	1.3379	1.2349	1.2328	1.4516
2018	0.6222	0.7327	0.8263	1.2340	1.1544
2017	0.6718	0.6115	0.5779	0.5628	0.6930
2016	0.4448	0.4790	0.4989	0.4912	0.7506
2015	0.3238	0.3603	0.3290	0.3654	0.4456
2014	0.3664	0.3262	0.3302	0.4086	0.4449
2013	0.3614	0.3377	0.2914	0.2789	0.3425
2012	0.5016	0.4111	0.3851	0.3110	0.3695
2011	0.4029	0.4774	0.4530	0.3819	0.3750
2010	0.3161	0.3568	0.3650	0.3875	0.3771

Table A.16: Corn Global Real Price Figures used in Model Phase 2

Year/TL	**Global Cor_P 1**	**Global Cor_P 2**	**Global Cor_P 3**	**Global Cor_P 4**	**Global Cor_P 5**
2022	8.4563	9.5775	8.9343	9.6609	9.3945
2021	3.2478	4.5058	4.0203	3.3837	6.2080
2020	1.7310	1.6678	1.7174	2.0714	2.5880
2019	1.5174	1.7566	1.8568	1.5777	1.6855
2018	1.1127	1.3370	1.2570	1.7182	1.5199
2017	1.0016	0.9817	1.0114	0.9026	0.9976
2016	0.7773	0.8492	0.7535	0.7208	0.9186
2015	0.7261	0.6966	0.8085	0.8271	0.7959
2014	0.7992	0.7559	0.5898	0.5286	0.6580
2013	0.9971	0.9285	0.8595	0.7027	0.6538
2012	0.8830	0.8437	10.771	10.507	0.9742
2011	0.8351	0.8727	0.8669	0.9444	0.8551
2010	0.4392	0.4418	0.4559	0.5672	0.6978