UNIVERSITY OF CALIFORNIA,
IRVINE

A framework for modeling the impacts of extreme events on the food and nutrition systems across the globe.

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Civil and Environmental Engineering

by

Alexandre Mathieu Pierre Martinez

Dissertation Committee:
Professor Amir AghaKouchak, Chair
Professor Steven Davis
Professor Kuolin Hsu

2021

DEDICATION

To

My friends and family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

These past five years of my life have been challenging and overwhelming, and I want to acknowledge the contributions of several individuals to this success.

First, I want to thank my committee members for their support, guidance, and time during these past five years. It has been incredibly helpful: Professor AghaKouchak, Professor Davis, and Professor Hsu, especially Dr. AghaKouchak, my primary mentor, for his support.

Then, I want to thank the administrative staff at UC Irvine for helping me to navigate UCI policies related to my Ph.D. requirements, travel expenses, et cetera: April Heath, Angie Zuniga, Jerry Martinez, UCI Police officers, the Office of Equal Opportunity and Diversity, and other staff members I have not interacted directly with.

Then, the University is not only about work, but it is also about connections, and these past five years would not have been possible without the support of my colleagues I met through work and can call friends now. I want to thank Felicia, Charlotte, and Yunxia from the same group, Negin and Mohammed from Sorooshian's lab, Lawrence, Clément, Antonis, and Janine from Efi Foufoula-Georgio's lab, and Paniz from Penn State University. Then, I want to thank my friends that I met at university and continuously supported me in the past five years: Tina, Shiva, Hossein, Elodie, Agata, Christian, and Sahar.

Then, I want to thank the French government for providing me with free quality education and health care from 2 years old in my hometown to my master's degree in Lille. I also want to thank Americans and emigrates in the US who are paying their taxes due to funding the National Science Foundation, the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration, and all public agencies are crucial to all doctoral students.

I also want to thank my uncle, Bernard, a farmer in France who helped me understand agriculture better, my brothers and sisters, who supported me and visited me these past years. Last but not least, I want to thank my parents, Antoine and Bernadette, who are in France, for their sacrifices during my childhood and their hard works.

VITA

Alexandre Martinez

Education

Doctor of philosophy	University of California, Irvine	2016-2021
	Civil and Environmental Engineering	
Master of Science	University of Texas at Austin	2011-2013
	Environmental and Water Resources Engineering	
Diplôme d'ingenieur	Ecole Centrale de Lille	2009-2011

Academic Employment

1. Research Assistant – University of California Irvine (October 2016 – Present)
    - Developed a framework to track the impact of extreme vents on nutrition supply. Developed several cloud-based tools to predict wheat's protein content and to map rice paddies.
2. Instructor of record – Texas Tech University (August 2013 – August 2016)
    - Developed courses content to teach fluid mechanic, hydrology, and hydraulics to undergrads from different majors.
3. Research Assistant – University of Texas at Austin (August 2011 – August 2013)
    - Modeling and laboratory experiments of mercury contamination in sediments.
4. Intern – CNRS/University de Lille II (January 2011 – June 2011)
    - Developed a tool to calculate the velocity fields in SPIV experiments.

Contribution to the field

Selected professional and academic meetings

1. AGU Fall Meeting (2020) – Talk.
2. AGU Fall Meeting (2019) – Poster
3. National Adaptation Forum, Madison, WI (2019) – Poster
4. Ocean Vision's meeting at Georgia Tech (2019) – Talk
5. AGU Fall Meeting (2018) – Poster

Selected public intervention and workshop organized

1. Drawdown program, Delaware County Institute of Science (2021) – Talk
2. Congressional meetings – Participated to and led several meetings with the California congressional staffs (Sen. Harris, Sen. Feinstein). Topics involved STEM Education, DEI in STEM, wildfire resilience, coastal resilience (2018 – now)
3. Climate change and water resources workshop, Truckee Earth Day (2019) – Funded by a grant from the AGU

Peer-review work (non-conference)

Journal papers – in progress

Martinez, AghaKouchak, Davis, et. al, "A framework for modeling the impacts of extreme events on the food systems across the globe." for Nature

Martinez, AghaKouchak, Davis, "A new macro-nutrition dataset MANU".

Teaching modules

Martinez, 2019, "Climate change analysis using MATLAB". Carleton College's Teaching computation in the Sciences Using MATLAB: Peer-Reviewed collection.

Hayatbini, Martinez, 2018, "Calculating and using Unit Hydrograph" Carleton College's Teaching Computation in the Sciences Using MATLAB: Peer-Reviewed collection

Professional affiliations

1. CUAHSI – HydroLean Fellow (2021)
2. CUAHSI – DEI Board member (2021)
3. AGU – Member of the Western Coalition / Wildfire Group (2020)
4. AGU – Fellow of the Voices of Science program (2020)
5. AGU – Board member of H3S section (2019)

Professional organizations

1. ASCE – Member 2013
2. AGU – Member 2016
3. AMS – Member 2018
4. AAAS – Member 2019

Awards

1. Broadcast Education Award. BEA (2021).
2. Inclusivity Award, University of California (2021).
3. New Face of Civil Engineering – Professional. ASCE (2020).
4. Henry Samueli Endowed Fellowship, University of California (2017).

ABSTRACT OF THE DISSERTATION

by

Alexandre Martinez

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Irvine, 2021

Professor Amir AghaKouchak, Chair

Extreme weather and climate variability carry out significant stress on the food production and distribution systems, already under pressure by increasing and shifting demands. In a globalized world economy with a highly interdependent network of food production and transmission systems, weather extremes cannot be regarded as local events anymore but rather as the triggering factors of a succession of downstream regional or even global impacts. Food and nutrition availability are national security issues, and the vulnerability of food distribution and nutrition supply systems are often assessed using metrics sensitive to local weather or climate events. Although informative, such metrics are often probabilistic and do not provide information on the local and nonlocal processes which drive changes in nutrition production, supply, and distribution. This study presents a framework for analyzing the interdependencies between countries' food systems, quantifying the nutritional content of the food produced, and quantifying the change in food and nutrition trade flows due to weather extremes or climate variabilities.

INTRODUCTION

Weather extremes pose significant risks to the global food system and the livelihood of many communities (Vogel & Meyer, 2018). Losses of crop yields are threatening farmers who do not have access to insurance products, communities whose primary source of income is agriculture, and countries at the edge of food insecurity. Roughly one-quarter of the global workforce is employed by the agricultural sector (International Labour Organization, 2019), and this number is even higher in developing economies (e.g., Chad, 75% in 2019) and major food exporters (e.g., India and Vietnam, respectively 43% and 37% in 2019, which are two leading producers and exporters of rice – a very labor-intense crop). Furthermore, crop insurance is widely available in middle to high-income countries but still rare in low-income countries (World Bank, 2018). Middle to high-income countries typically use named-peril (such as hails storm) or multi-peril crop insurance (MPCI), while low to middle-high income countries are experimenting with index-based crop insurance, such as cumulative rainfall or crop yields aggregated in a geographical area. While farmers are aware of their production risks and can use their knowledge and experiences for their crop management, changing climate resulting in new frequency and strength of weather extremes will likely disturb their farming practices and create additional production risks.

In addition to weather extremes and changing climate, rising economies and demographic changes will add more stress to the food production systems. Calories and protein demands are expected to increase respectively by 60% and 110% by 2050 (Alexandratos & Bruinsma, 2012). Eight hundred million people are already chronically suffering from hunger (World Health Organization, 2016) and, despite progress in some

regions of the world, the hungry population is still growing (Godfray, et al., 2010), (FAO, 2012). Increasing competition for non-food uses of crops (e.g., biofuels) and rising living standards (e.g., increased meat consumption) are both continuously magnifying food insecurity.

Nonlocal changes in food supply

Internationalization and globalization have drastically transformed the global food system within the past decades (Wilkinson, 2009). Major industrialized food producers continuously strengthen their production and exportation capacities and dominate the global food trade network. On the one hand, a more globalized food system can buffer against local food crisis, and the rate of people dying due to a famine decreased globally from 50 per 100,000 persons in the 1920s to 1960s to 2 per 100,000 persons in the 1990s and 0.5 per 100,000 persons in the past decades (2010s) (Rulli & D'odorico, 2013). But on the other hand, a more globalized food system exposes countries to production shocks happening in other parts of the trade network. Moreover, the gains in dietary energy supply (i.e., calories) are uneven and accompanied by a few drawbacks. For instance, for many countries, the increase of the dietary energy supply thanks to trade resulted in a less diverse food being consumed (Khoury, et al., 2014), and high-value staple crops have replaced traditional crops, decreasing the consumption of micro-nutrient (Webb, 2009). Overall, the global food system moved toward a concentration of a few commodities (cereals) produced by a few major food-producing countries. Table 1 shows the share of the top producers from the most consumed commodities. While wheat concentration is relatively low (Five countries are growing half of the global wheat), other commodities are

much more concentrated: three countries are growing half of the global rice, and two countries are growing half of the global corn.

Wheat		Rice		Barley		Maize	
Country	**Share**	**Country**	**Share**	**Country**	**Share**	**Country**	**Share**
China	17.6%	China	28.4%	Russia	11.9%	United States	33.7%
India	12.9%	India	22.0%	France	7.9%	China	22.6%
Russia	8.5%	Indonesia	8.0%	Germany	7.5%	Brazil	7.5%
USA	7.8%	Bangladesh	7.1%	Australia	6.3%	Argentina	3.5%
France	5.2%	Viet Nam	5.9%	Ukraine	5.9%	Indonesia	2.2%
Total	**52%**	**Total**	**76%**	**Total**	**40%**	**Total**	**70%**

Table 1 – Major players of cereal market (2010-2019) : Top five producers of cereals and their shares to the global production of wheat, rice, and corn during the 2010-2019 period. Data source: FAOSTAT

There is also a concentration of exported food toward cereals, as shown in Figure 1, revealing the increase of wheat, rice, and corn in calories and proteins traded from 2000 (represented by a diamond) to 2010 (represented by a circle). This concentration toward a few commodities and few producers has created an unprecedented situation for the nutrition market (Garlaschelli & Loffredo, 2005), (Fagiolo, Reyes, & Schiavo, 2010), (Ercsey-Ravasz, Toroczkai, Lakner, & Baranyi, 2012) in which extreme events resulting in crop failures in one country could cause or heighten a short-term food shortage in other countries. Food security cannot solely consider the resilience of domestic food systems but shall also include the resilience of the major producers and the resilience of a country's trade network.

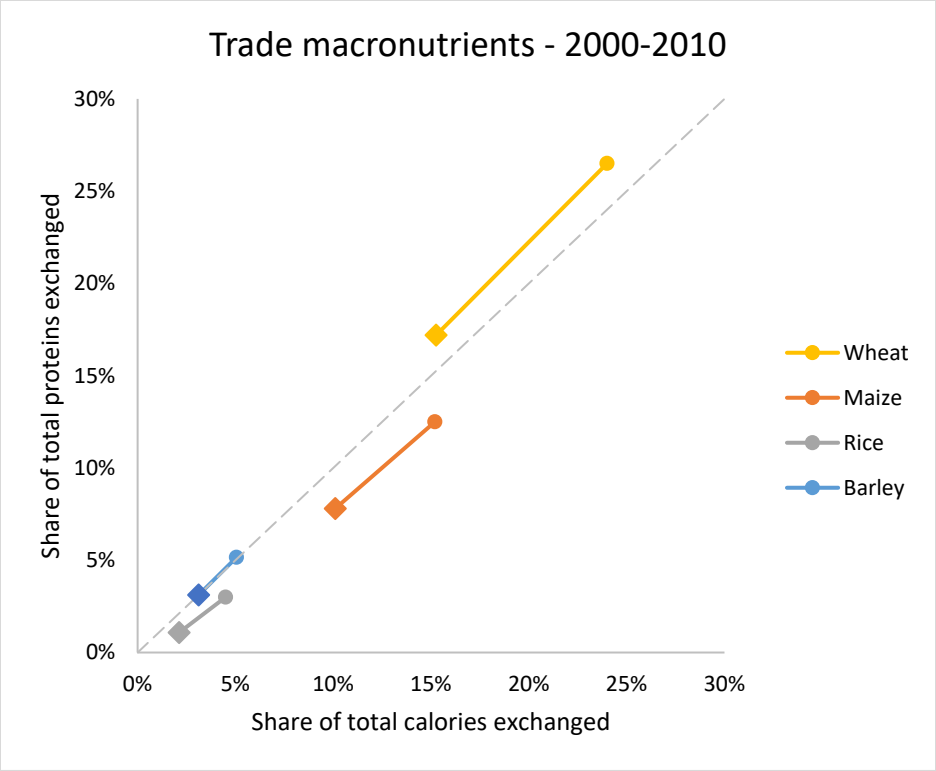

Figure 1 – Concentration of cereals in the food basket : Evolution of the shares of wheat, corn, rice, and barley in the total calories and proteins traded. The year 2000 is indicated by a diamond (◆) and the year 2010 by a circle (●).

Nonlocal change can indeed impact local food security. For instance, analysis of the food supply from FAOSTAT reveals that some countries' most significant decreases in food supply were caused by nonlocal reductions in supply (i.e., trade) rather than local decreases in supply (i.e., production or storage). For instance, Egypt and Algeria experienced four years with a significant decrease in wheat supply from 1990 to 2010. Figure 2 shows that two and three out of four were nonlocal respectively for Algeria and Egypt.

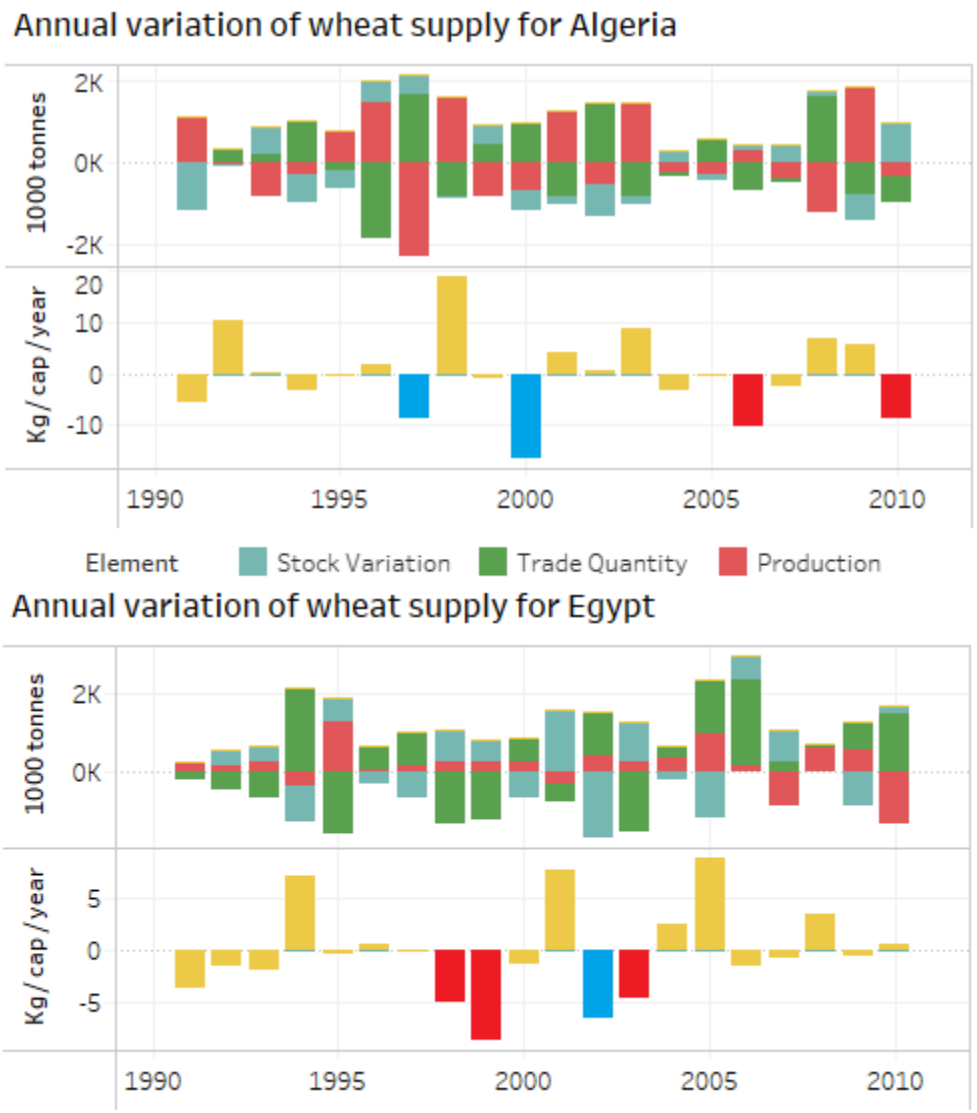

Figure 2 – Local and nonlocal variations of wheat supply : example in Egypt and Algeria. In the lower part of each chart, the red color indicates a nonlocal decrease (due to trade), and the blue color indicates a local decrease (due to production/storage).

In the following sections and chapters, we will use several indices defined as follow:

- The Importation Dependency Ratio (IDR) indicates the dependency of a country on importation relative to the total domestic consumption, and it is calculating such as (FAO, 2020):

$$IDR = \frac{Import}{Production + Import - Export}$$

- The Dietary Energy Supply is defined as the average country-scale supply of calories (kCal per capita per day) from all commodities consumed (FAO, 2020):

$$DES[cal/cap.day] = \sum_{Food\ k} S_k[kg/cap.day] \cdot C_k[cal/kg]$$

- The Minimum Dietary Energy Requirement is the minimum energy (kCal per capita per day) required for a population to maintain acceptable body weight for attained height. It is calculated by the World Health Organization (WHO) every five years (FAO/WHO/UNU, 2001).
- The Herfindahl-Hirschman Index (HHI) is a standard measure of market concentration and is primarily used to determine market competitiveness after companies' mergers and acquisitions. It is expressed as (Rhoades, 1993):

$$HHI = \Sigma s_i^2$$

Where Si is the share of the element i expressed in percentage. Its values range from 1 (No concentration) to 10,000 (High concentration). The value of 10,000 indicates one company having 100% of the market and all other companies 0%.

For ease of use, we define here a unit diversity score derived from the HHI such as:

$$D = 1 - \frac{HHI}{10{,}000}$$

Which is now ranging from 0 (Low diversity) to 1 (High diversity).

For supply diversity, S_i represents the share of the country i in the total importation of nutrition (calories or protein) by the country of interests j such as:

$$HHI_j^{supply} = \sum_{i=1}^{N} \left(\left(\frac{T_{i \to j}}{\sum_N T_{k \to j}} \right)^2 \right)$$

Where $T_{k \to j}$ is the total nutrition (calories or proteins) traded from the country k to the country j, and N is the total number of countries.

For diet diversity, S_i represents the share of the commodity i in the total consumption of nutrition (calories or protein) by the country of interests j such as:

$$HHI_j^{diet} = \sum_{i=1}^{N} \left(\left(\frac{f_i}{\sum_N f_k} \right)^2 \right)$$

f_k is the consumption of nutrition (calories or proteins) from the commodity k in the country j, and N is the total number of commodities.

Figure 3 depicts global trade and dependencies between countries (figure generated based on the proposed model discussed in the following section). A country is vulnerable when its average Dietary Energy Supply (DES) is not significantly higher than the WHO's Minimum

Dietary Energy Requirements (MDER). It indicates that even a minor disturbance in food delivery/production/import can lead to a nutrition deficit.

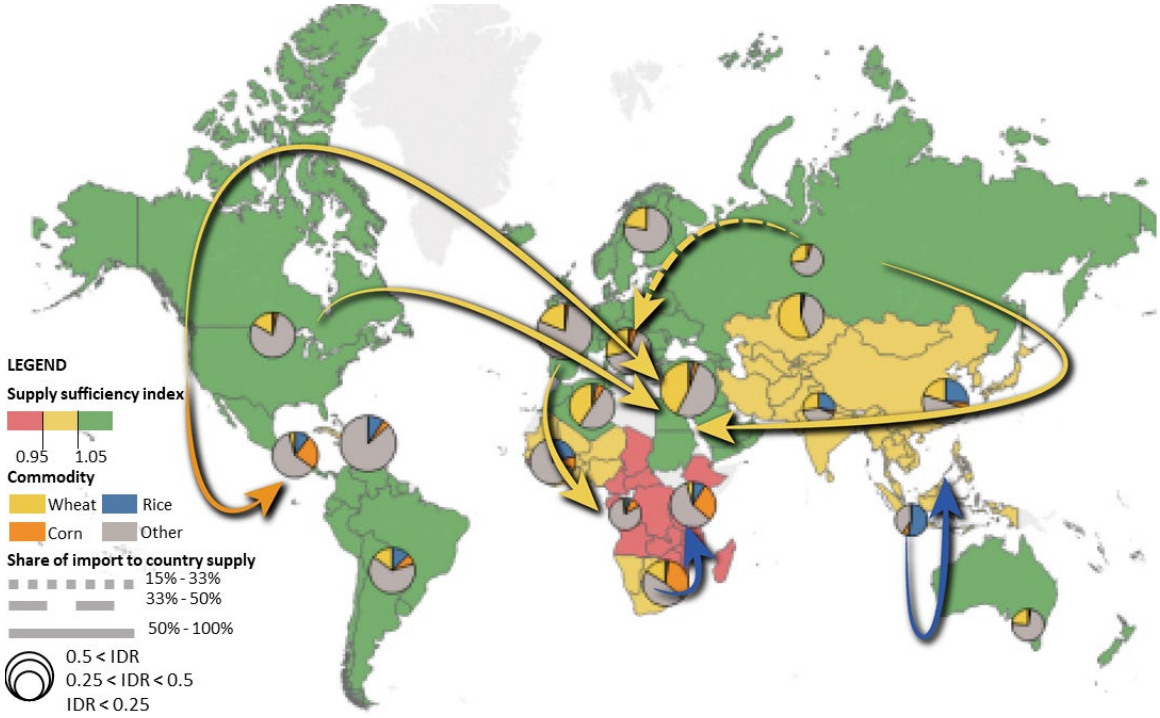

Figure 3 – Global food security situation during the 2000-2010 period. The arrows correspond to the main food flows between regions, the background color symbolizes the dependence on importations, and the pie charts represent the main commodity providing dietary energy.

The pie charts represent the share of wheat, rice, maize, or other commodities in the regional calory supply (North American, Central America, South America, Caribbean, Western Europe, Northern Europe, Southern Europe, Russia, North Africa, West Africa, South Africa, East Africa, Middle East, Australian, Southeast Asia, Southwest Asia). The size of the pie chart represents the dependency of the region on importation (IDR calculated at the regional scale). The primary trade flows between regions are represented by arrows,

whose color indicates the commodity, and the style (dots, dashes, straight line) indicates the share of importation in the regional calory supply.

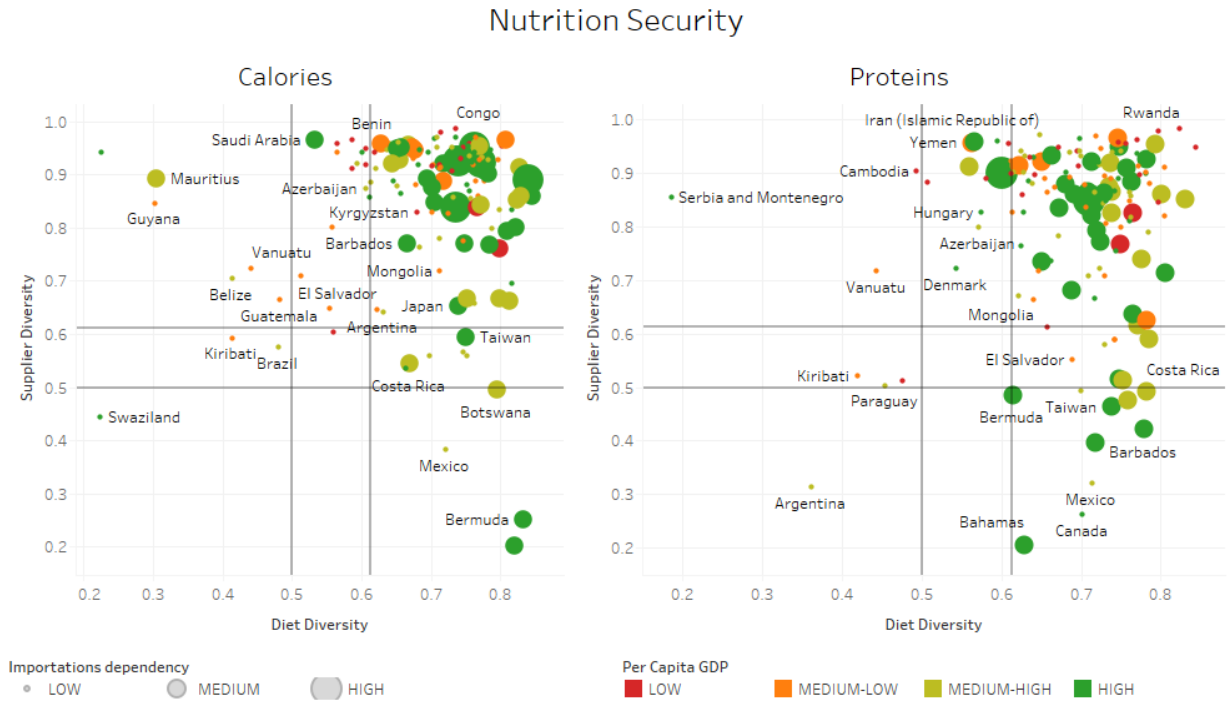

Figure 4 – Nutrition diversity scores during the 2000-2010 period. Supplier diversity (y-axis) and diet diversity (x-axis) for calory (Left panel) and protein (Right panel). The color symbolizes the nation's wealth, and the circle's size is proportional to the country's dependence on importation.

Figure 4 shows the diversity of the food consumed and imported for each country from 2000 to 2010. The axes represent the coefficient of diversity, calculated using complementary of the Herfindahl-Hirschman Index (HHI) normalized to the unit. We used a threshold of 0.5 and 0.62 to classify the diversity score between low, medium, and high diversity. A score higher than 0.62 is considered here as a high-diversity score. The threshold of 0.62 is derived from a commonly accepted threshold defined by the US

Department of Justice to calculate market concentration and authorize or deny mergers. The color represents the per capita GDP categorized according to the 2010 United Nations' thresholds, and the size of the circle depicts the dependency on importations for food supply. Overall, there is much less diversity of food categories consumed or countries' suppliers for proteins than calories. The countries with low diversity are the most at risk, while the countries with high dependence to import and high per capita GDP are most likely to propagate production or trade shocks to other countries with lower buying powers.

The knock-on effects of weather extremes and climate variations on food production

All changes in climatic variables such as minimum temperature, maximum temperature, and precipitation patterns are expected to drastically disturb food production, but these risks are not well-perceived yet. A recent study from Yale University showed that most Americans believe that climate change is happening (70% national average) and will harm plants and animals (70% national average) (Howe, Mildenberger, Marlon, & Leiserowitz, 2015). However, much less think that climate change will harm the USA (58% national average) or them personally (41% national average), as shown in Figure 5. This difference comes from the difficulties of tracking the impacts of climate change and extreme events outside of the geographical area where it happens. For instance, the mechanisms by which extreme events impact farmers are complex and not all fully understood yet. Considering that 80% of the US corn exports transit by the Mississippi River to be shipped worldwide from Louisiana harbors, the farmers' revenues

in the Midwest can be potentially impacted by sea-level rise, flooding, and hurricanes in the Gulf of Mexico. More information can be found in Appendix 0.1.

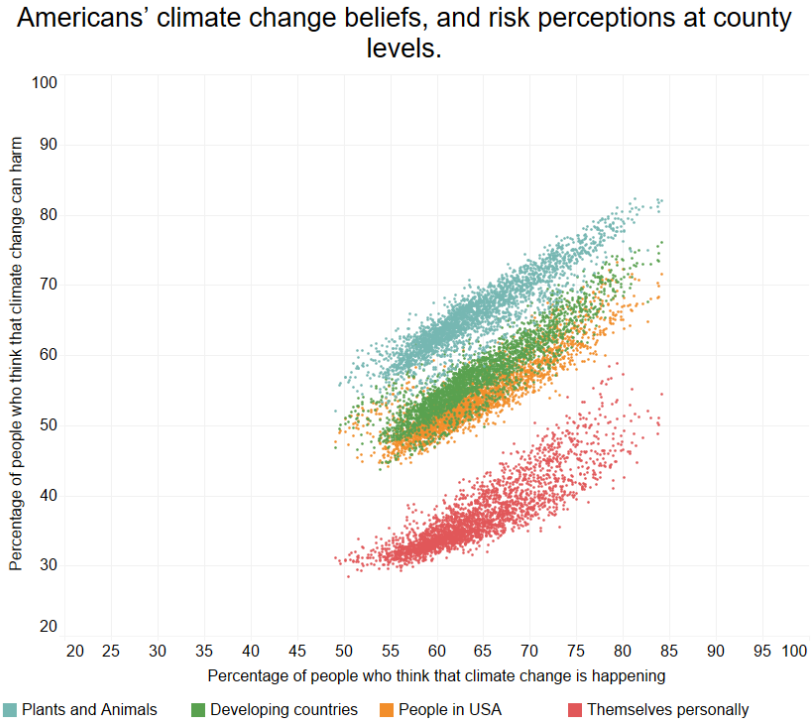

Figure 5 – Perception of climate change beliefs and risk perceptions in the USA in 2010. Each point a US county's perception of climate change threat to plants and animals (blue), to developing countries (green), to people in the USA (orange), and to themselves (red). Data are from the Yale Climate Change Communication (Howe, Mildenberger, Marlon, & Leiserowitz, 2015).

Therefore, it is critical to understand how local extreme events in major food exporters can propagate through and interact with the global food systems. Chapter 1 proposes a framework to answer this question.

Changes in extreme weather will make it more challenging to forecast food production, prices, and vulnerabilities, especially across large distances. In addition, we can

expect more production shocks will systematically propagate to other countries and have magnified impacts in vulnerable regions. Food prices are based on two factors: the quantity offered and the demand, and the production quality, which will remain highly uncertain in the following decades. Globalization of the food system is also expected to increase because of the leading food producers' intense industrialization and production capacity.

Although yields have consistently increased globally within the past decades, substantial variations and significant losses of yields are still happening often. We analyzed yields in cereals (Appendix 0.3. for more details) for all countries to quantify the number of years exhibiting significant drops in yields. In Figure 6, countries are ranked by their total production of wheat (we used 2019's data from the FAOSTAT, left y-axis on the figure), and grouped by 50%, 90%, 95%, and 99.5% of the total production: dark green background indicates 50% of the total production, lighter green indicates 90%, et cetera. The number of yield decreases is displayed by magnitude (5%, 10%, and 20% or more decrease) on the left y-axis. Generally, the top five producers (China, India, Russia, United States, and France) experience less substantial yield variations (variation over 20% decrease) than other countries.

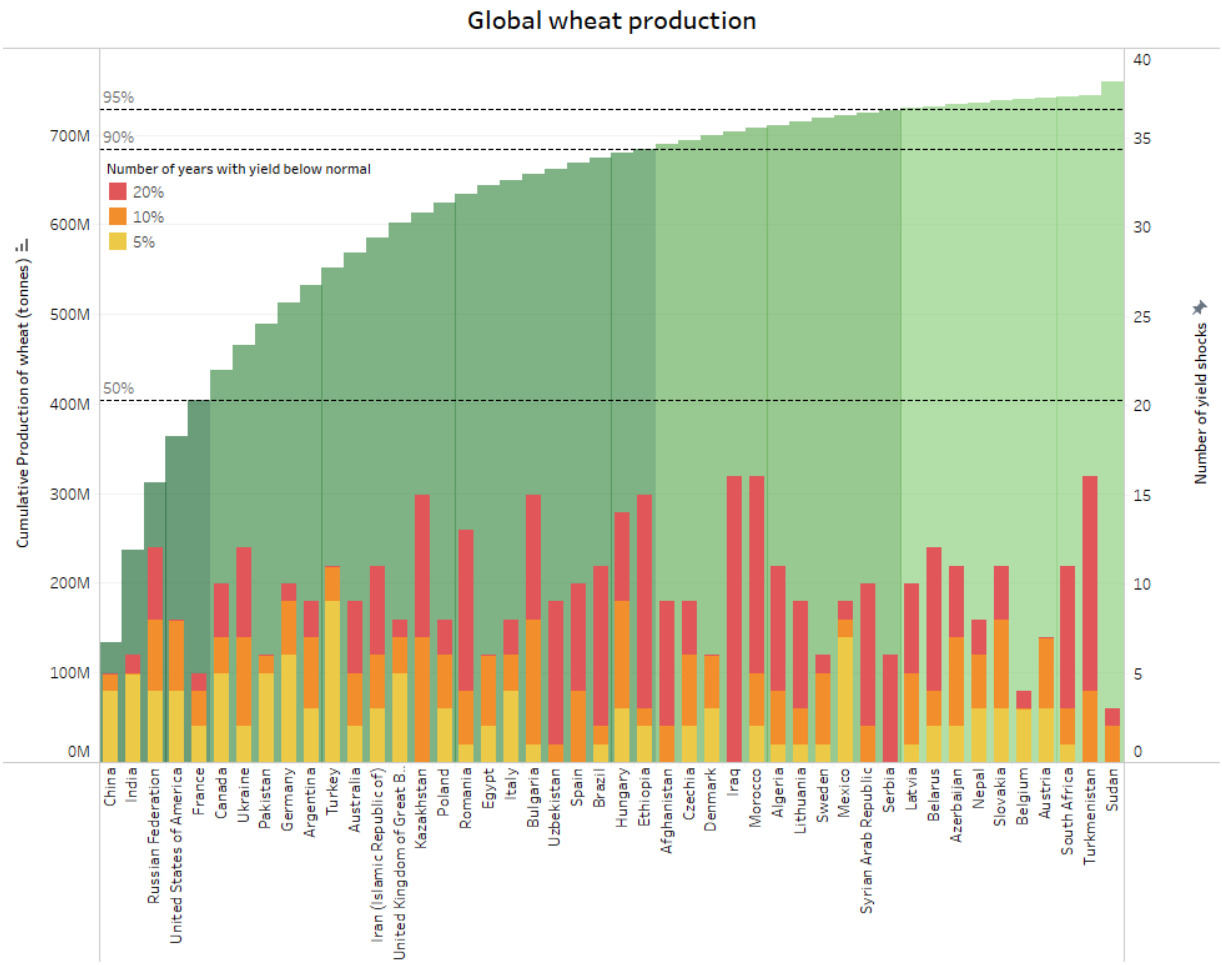

Figure 6 – Global wheat yield shocks by magnitude for the top producers (95% of the global production).

Although relatively moderate and mainly under 5% lower than the average, such variations can be due to severe weather conditions. A similar analysis has been performed for maize, as seen in Figure 7. Similar to wheat, patterns can be observed, with lower-in-magnitude yield shocks for the first top producers (the United States and China) and higher for the other producers. However, Brazil and Argentina, two other primary producers, have much more large yield variation than the United States and China.

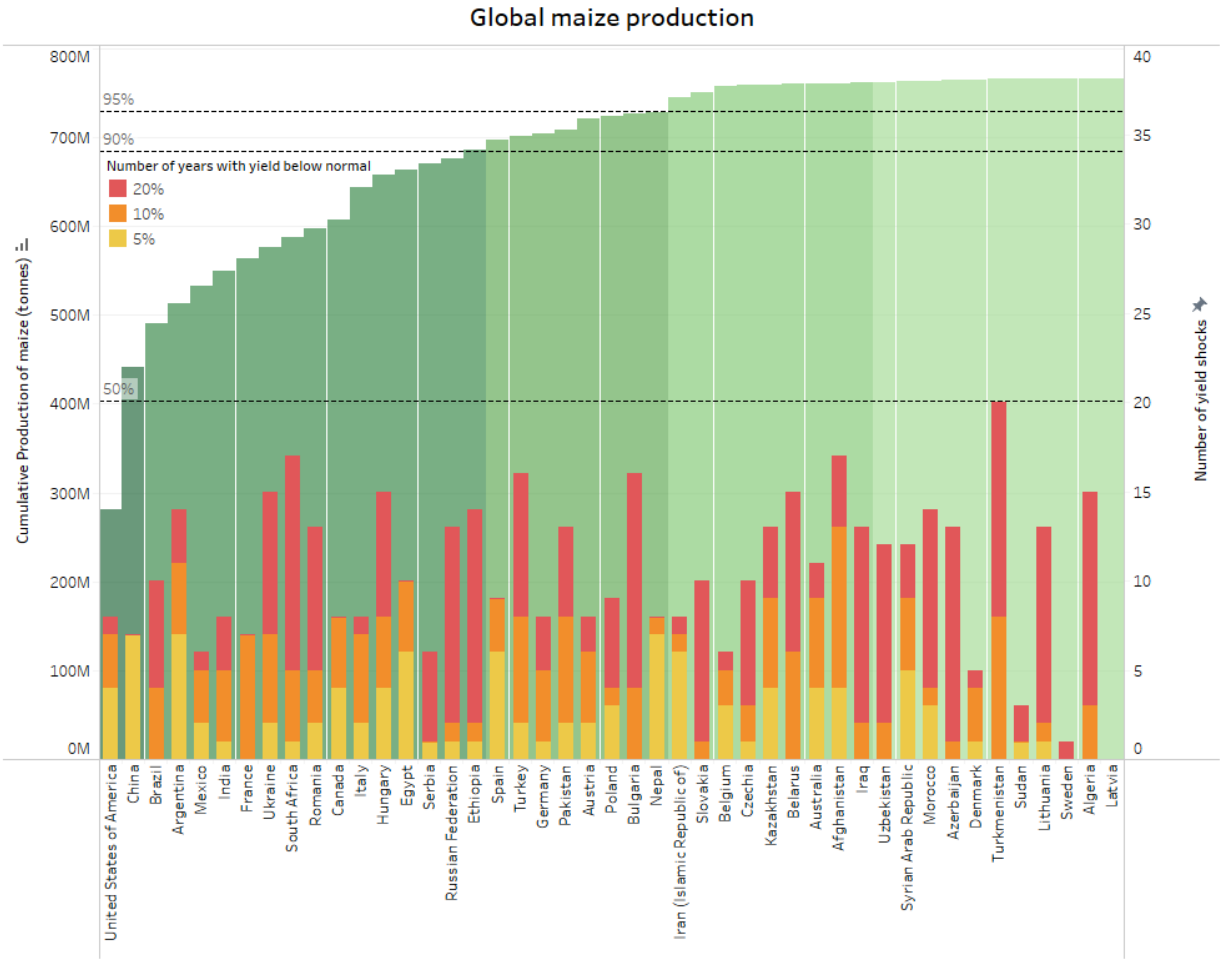

Figure 7 – Global maize yield shocks by magnitude for the top producers (95% of the global production).

Climatic variations or extreme events are one of the leading causes of losses of yields. Figure 8 summarizes the natural disasters from 1990 to 2019 recorded in the Emergency Events Database from The Centre for Research on the Epidemiology of Disasters (CRED) within the Université Catholique de Louvain (UCLouvain) (CRED, 2019). For instance, United States experienced seven major droughts within the past 15 years

while China experienced 18 significant droughts, which added stress on water resources, irrigation, and cereals production. Flood is also a major concern in China, with more than 30 major events in China, the USA, and India that could also result in significant disruption of food production, such as delays in planting operations or destruction of crops of harvest stored.

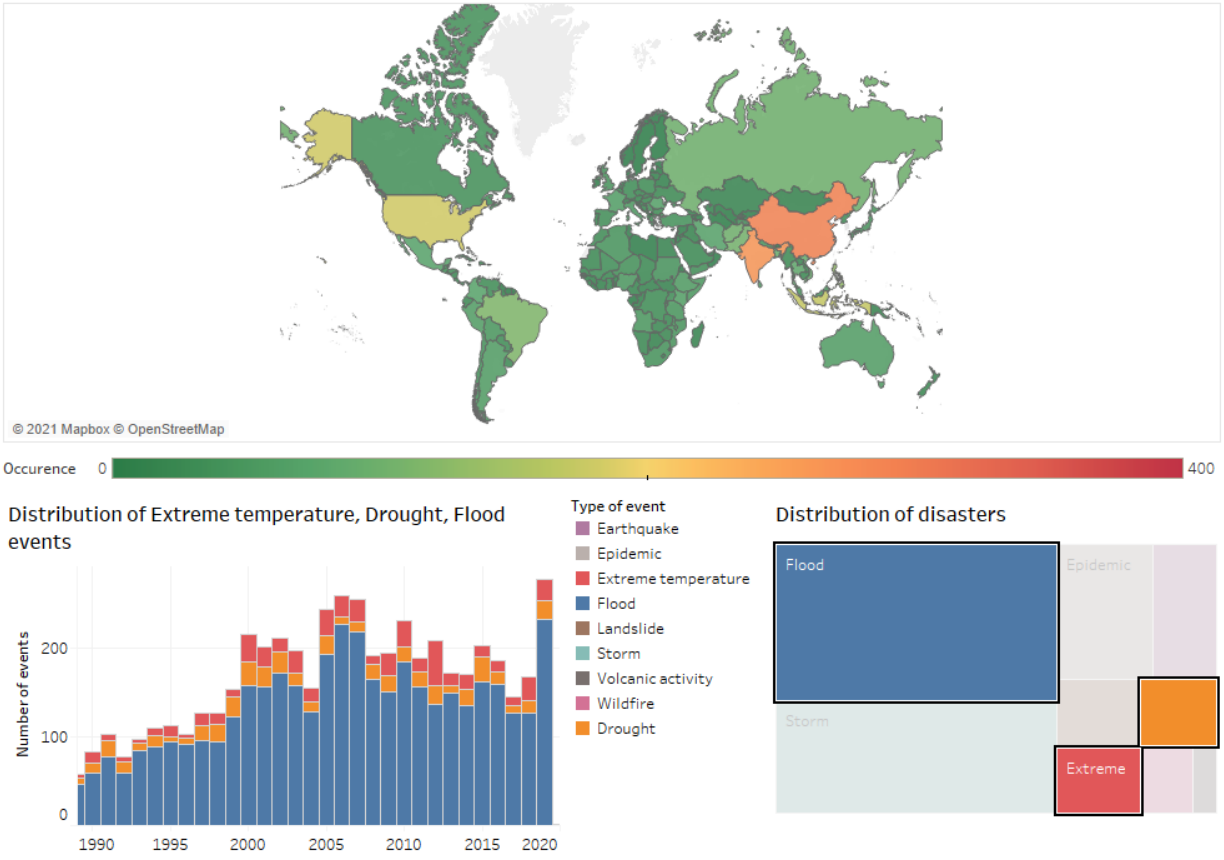

Figure 8 – Global distribution of natural disasters by category from 1990 to 2019. The top map displays the occurrence of all combined major floods, droughts, and heatwave events, the bottom left chart indicates the temporal evolutions of it.

In the United States only, major droughts, heatwaves, freezes, floods, or storms affecting crops are frequent (Smith & Katz, 2013). Furthermore, climate change is expected

to deeply affect food security at both local and global levels (Reilly & Kane, 1992), (Rosenzweig & Parry, 1994). First, the number of droughts happening in some parts of the world is expected to increase. Then, droughts themselves are also expected to change in duration, intensity, and temperature (Chiang et al. 2018; Chiang et al. 2021). It has been found that droughts in the United States are getting warmer faster than the rest of the climate (Chiang et al. 2018). Furthermore, significant increases in concurrent droughts and heatwaves have also been observed in the United States (Mazdiyasni & AghaKouchak, 2015), resulting in additional stress on crops.

Moreover, precipitation is expected to slightly increase in the northern states, the main center of cereals. Heavy downpours are expected to drastically increase, resulting in more intense flooding on the crop areas. Spring flooding typically means livestock displacement, winter wheat destruction, storage bin destruction, and delays in planting corn, soybean, or spring wheat. If the flooding is long enough, there will not be enough time for the farmers to seeds their crops for this year. Summer heavy precipitation or floods can also happen and would result in the delay of the harvest or the destruction of the crops. Change of flood intensity is also adding pressure to the existing levees networks suffering from low maintenance and a higher risk of levee failure. Because most people heavily rely on government infrastructures to protect them from floods, they often store their grains in areas that can be easily flooded in case of levee failure. During Spring 2019, flooding in the Midwest destroyed most of the storage bins that were on site. Low cereals prices, lower demand due to the US-China trade war pushed the farmers to store their harvests on-site in 2018 more than usual, and these have been heavily damaged or destroyed during the flooding events. There are some opportunities to delineate flooded areas using SAR data.

SAR data can provide valuable long-term information on flood extend without having to run hydrological modeling. On a different scale, extreme floods in the Mississippi river often disrupt barge operations, with the direct consequences of farmers having to store their grains on site. Tariffs with China also created a situation in which farmers had to store their grain on-site, on unprotected uninsured sites until the exportation can happen at a better price, resulting in increased grain storage in at-risk areas. The 2019 flood impact is an example of the destruction of the previous harvest stored more than usual on the field (due to tariff) and delayed and destruction of future crop production.

Surface temperature (ST) is known to affect crop quality and quantity. An overall increase in temperature may have positive impacts on cereal production (Adams, et al., 1990), (Parry, Rosenzweig, & Livermore, 2005), although recent researches have suggested that this is more complex than it seems and that higher diurnal temperature might also lead to lower yield (Lobell, Changes in diurnal temperature range and national cereal yields, 2007). Furthermore, winter or early spring heatwaves can be devastating from agriculture from subsequent frosts. For cereals, heat stress during the grain-filling phase affects the grain protein contents (Wardlaw, Blumenthal, Larroque, & Wrigley, 2002). Predicting the protein content using a physical-based model or remotely sensed information is still highly hazardous and remains one of the main challenges (Hansen, Jorgensen, & Thomsen, 2002). Higher temperature during the grain filling phase results in a higher protein accumulation rate in the grain up to a threshold (typically 30^{o}C). Chronic high temperature or heat shocks, however, doesn't affect the rate of protein accumulation significantly but negatively impact the duration of the grain-filling phase and, therefore, the overall protein content (Randall & Moss, 1990), (Panozzo & Eagles, Cultivar and

environmental effects on quality characters in wheat. II. Protein, 2000). Wheat heat shock occurs above 32°C, while the cardinal temperature range of other crops is 40°C – 42°C (Hatfield & Prueger, 2015). With global mean temperature projected to rise from 1.4°C-7.8°C, with higher temperature rise in some parts of the world, we can expect to have more and more heat shock on cereals in some parts of the world.

Elevated Atmospheric CO2 concentration has been consistently reported to increase the crop yields due to higher photosynthesis activity on the one hand (Kimball & Idso, 1983), (Ainsworth & Long, 2004), (Lobell & Gourdji, The influence of climate change on global crop productivity, 2012) but to decrease the leaf-N concentration and the protein content in wheat on the other hand (Benzinger, Kyle, & Blumenthal, 1998), (Fangmeier, et al., 1999), (Hogy & Fangmeier, 2008), (Panozzo, et al., 2014). Open-air experiments conducted in Australia and Germany have reported a decrease of respectively 3.7% and 7.4% of the protein content of wheat under high CO2 concentration (550ppm vs. 380ppm) (Panozzo, et al., 2014). Generally, higher carbon concentration is expected to make the food produced less nutritious (Myers, et al., 2014). C3-plants, such as wheat, rice, two primary staple foods on the planet, are typically sensitive to atmospheric carbon concentration, while C4 plants such as maize and sorghum have not been found to be. C3-plants, such as wheat, have been projected to have up to 6% less protein by 2050 under higher atmospheric carbon concentration (Myers, et al., 2014), typically of 200ppm, while C4 plants, such as maize, can minimize photorespiration and the impact of higher CO2 concentration. For 25% of the population, mainly located in North Africa and the Middle East, more than 50% of the proteins come from C3-plants, as shown in Figure 9. Middle

East, Southern Asia, and Southeastern Asia rely the most on C3-plants for their protein supply.

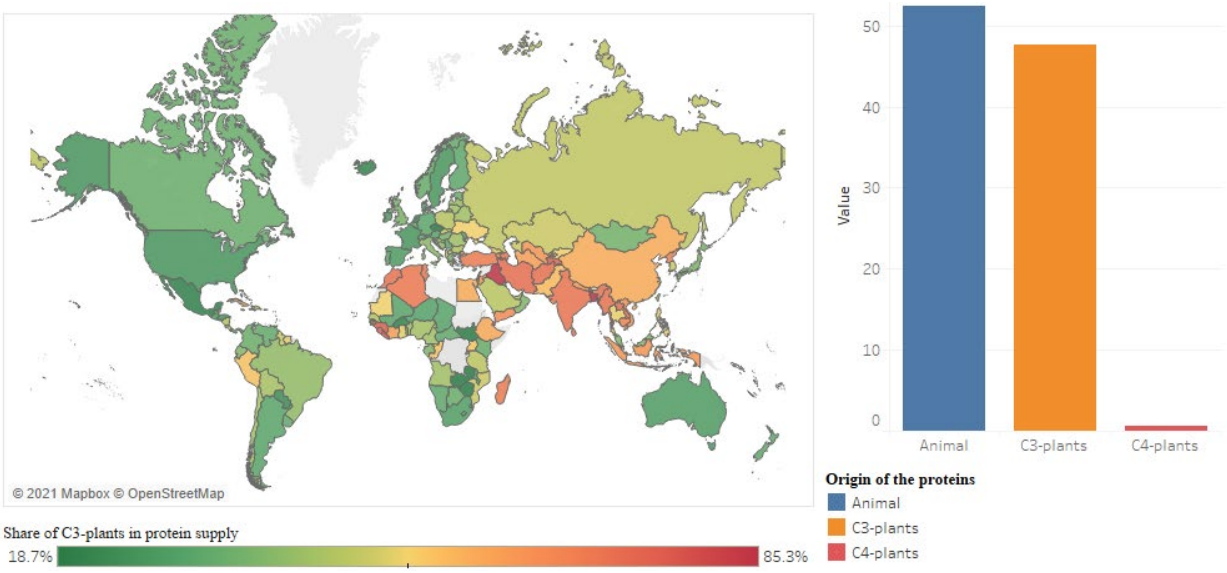

Figure 9 – Source of protein consumption. Left: A map of the share of C3-plants in the domestic protein supply and right: Global average of the source of protein.

Furthermore, most of the crop losses in the USA due to air pollutants have been attributed to ambient ozone (O3), alone or along with acid rain precursors (Tong, Mathur, Schere, Kang, & Yu, 2007). Other air pollutants that affect crops are particle matter (PM), acidic compounds, ammonia, nitrogen and phosphorus (nutrients), black carbon (BC). Air pollutions affect crops through different mechanics: direct (reduction of incoming solar radiation, photosynthetically active radiation PAR; deposition of pollutants on plants, e.g., aerosols; acidification or acid rain, damage to plants due to ozone) or indirect (change of climate conditions, e.g., temperature, humidity, clouds, extreme events or change the hydrological cycle and water availability). "UVs are not supposed to decrease crop production, it might be the opposite actually depending on the dose, and most of the

research are lab experiments that can't be directly applied to the field." (Wargent & Jordan, 2013). Summer is more sensitive to O3, while winter is more sensitive to SO2 and NO2 combination (Agrawal, Singh, Rajput, Marshall, & Bell, 2003). The higher temperature is decreasing crop production, and ozone pollution is different depending on the crop. A study (Tai, Martin, & Heald, 2014) has been conducted on wheat, rice, maize, and soybean. It used the CESM (Community Earth System Model) to project future temperature and ozone concentration under different climate change scenarios RCP4.5 and RCP 8.5, and it has been found that ozone concentration is expected to decrease in some regions (US, Europe) and thus increase production, but ozone is increasing in south Asia and decreasing production. The global effect of ozone is an increase of production under the RCP scenario 4.5 but a decrease under the RCP scenario 8.5. PM exposure and transport of toxic particles through PM can also create vegetation stress (Grantz, Garner, & Johnson, 2003). For instance, PM can absorb infrared, diminish the surface solar visible radiation by 7-18% leading to modest but significant changes in crop yields (Chameides, et al., 1999), (Grantz, Garner, & Johnson, 2003). Surface ozone (O3) is a potent phytotoxic air pollutant that reduces the productivity of crops (foliar injury, accelerated senescence, decrease in growth, altered metabolism, reduced ability to sequester carbon, and reduced crop yield). Roughly, if ozone concentration doubles (from 40 to 80), crop yield is divided by 2 (from 55 to 35 bu/acres). South Asia is the most polluted region in the world. There are some seasonal variations of ozone concentration. For example, in South Asia, rice and Soybeans are grown by the end of the high ozone seasons (April), but wheat is grown during the high season. O3 concentrations have a strong seasonal and regional pattern. Most of the current studies focus on the impact on the current cultivated areas (not areas projected to become

cultivated by 2050+) and only economic losses (not checking the change in water intensity or energy intensity). In Europe, peaks of ozone are expected to be fewer thanks to actions reducing local pollution. However, the background ozone concentration is projected to increases due to long-range transboundary transport from other parts of the world. (Sutton, Amann, Emberson, & Dentener, 2015). Fertilizers are less efficient with a high ozone concentration. Plants are also less drought-tolerant (decrease in root biomass, limiting access to soil water). Irrigated crops are more sensitive to ozone than rain-fed crops (Teixeira, et al., 2011). Adaptation methods include shifting crop calendars by changing sowing dates, applying irrigation, and using crop varieties with different growth cycles. This will reduce the losses due to ozone (plant growth during the "lower" ozone season; however it might require more irrigation due to lack of rain in the dry season. South Asian countries are the most at risk. 40% of the damages are predicted to happen in China and India (Van Dingenen, et al., 2009). USA ranks third with approximately 25% of the global losses but ranks first for soybean losses (estimated to be $2bn) and second for corn ECL (estimated to be $0.6bn). These losses are expected to increase with time (More in appendix). Wheat, rice, and soybeans are the most sensitive to ozone pollution and their yield and nutritional quality can be affected. (McCool, Musselman, & Teso, 1986), (Feng & Kobayashi, 2009), (Vlachokostas, et al., 2010). In the year 2000, it is estimated (Avnery, Mauzerall, Liu, & Horowitz, 2001) that the losses were 8.4%-14% for soybean, 3.9%-15% for wheat, and 2.2%-5.5% for maize, for a total loss of 79-121 million metric tons (about $11-18 $b). For economic losses, some prices are regulated in Europe so if supply decreases, the price might not change, and thus economic losses are directly related to yield losses.

Overall, extreme events, air pollutants, and rising CO2 creating more variable weather are all undermining the quality (e.g. cereal's grain size, protein content) and the quantity of the food produced (Myers, et al., 2014), (Myers, Wessells, Kloog, Zanobetti, & Schwartz, 2015), (Medek, Schwartz, & MyersqS., 2017), (Zhu, et al., 2018). It is therefore important to be able to better understand the nutritional value of the food produced, traded, and consumed.

Being able to track the nutritional value of the food produced, traded, and consumed is therefore very a key aspect of food security and will be presented in chapter 2.

Importance of the nutritional content of the food produced.

Current research efforts on impacts of climate change or extreme events on food security are mainly restricted to change in crop yield and the consequent change in calories supply. Figure 10 highlights the focus of recent publications on crop yield for food security while not considering crop's quality as much as it deserves. We looked for journal publications on the Web of Science using the titles keyword, we remove conference proceedings, review papers, and kept only papers that have developed and tested a model with actual data.

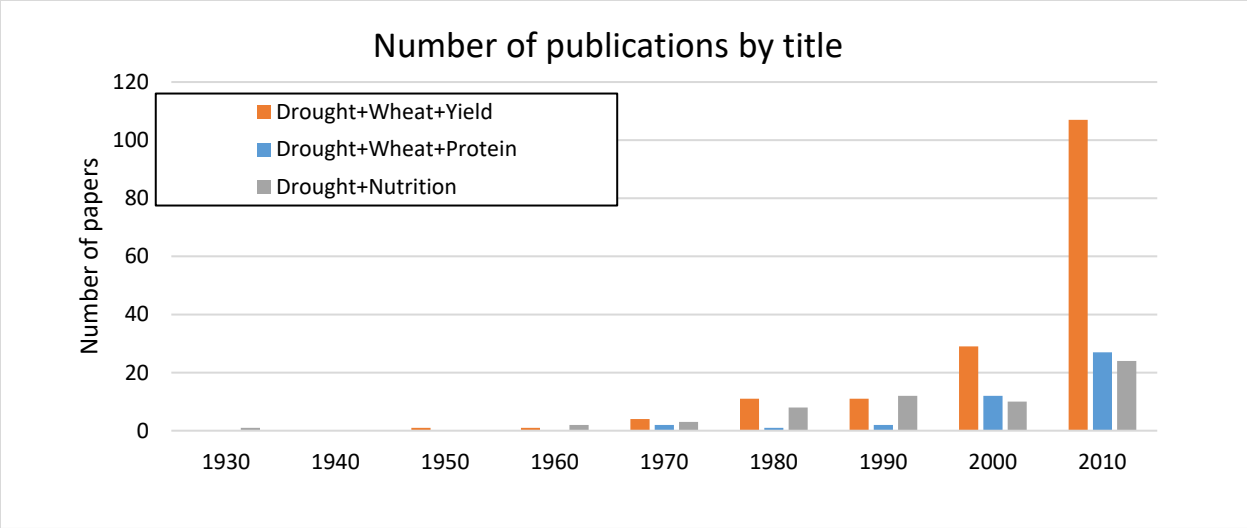

Figure 10 – Publications categorized by keywords in the title for the past 8 decades.

Changes in other dimensions of micronutrition (such as fat and protein) recently got more attention, but unfortunately, these aspects received much less attention than they deserve. Variation of protein content in cereals, mainly due to the plantar used and local growing conditions, is not negligible. For instance, the Hard Red Winter Wheat (HRW), one of the most grown species in the United States, has a median protein content ranging from 11.5 to 12.0 g/100g-wheat and a full range going from 9.0 to 15.0 g/100g-wheat as shown in Figure 11 (data from the Federal Grain Inspection Service (USDA-FGIS)).

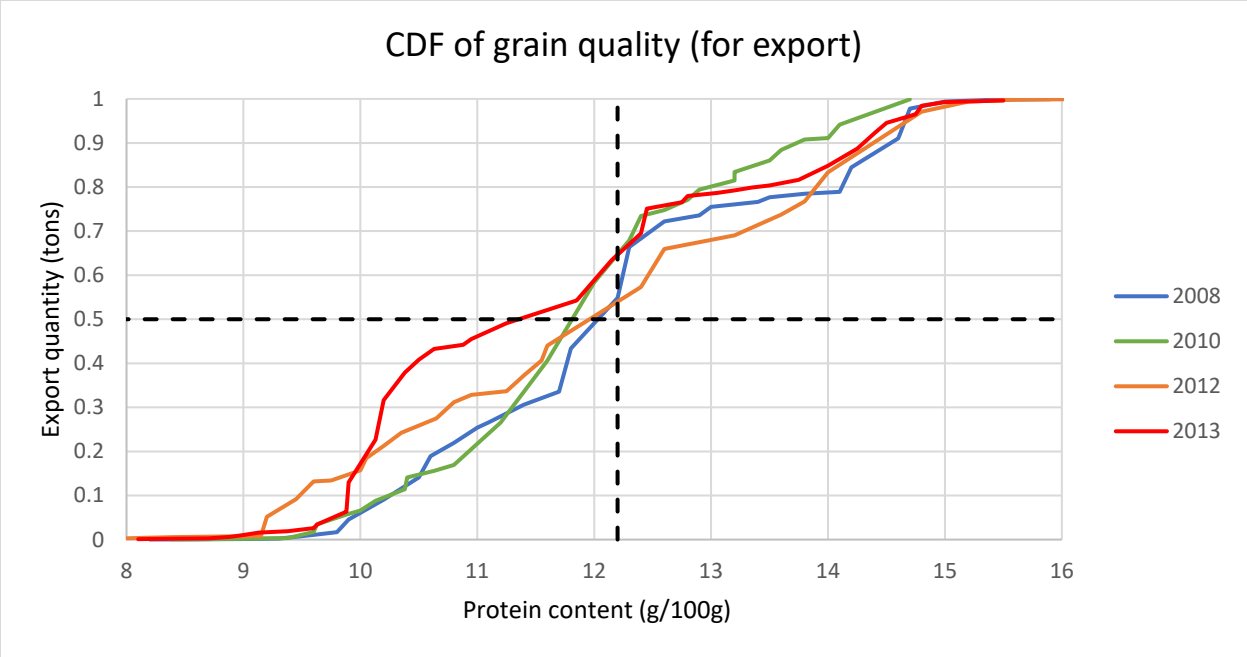

Figure 11 – Exported Hard Red Winter wheat's protein content : cumulative distribution Function. The horizontal dash line represents the median and the vertical represents the standard value of 12.5% protein content.

A small change in crop quality is unlikely to affect both primary producers and importers. In general, a fraction of the population in countries with a less diverse diet and a small surplus of protein is often impacted by protein deficit. We estimated that a change of 20% of the protein content of the U.S. wheat (from 12g/100g down to 10g/100g) would result in a change of 40% of protein supply in several countries (such as Belize), that relies heavily on the importation of U.S. wheat. We estimated that a 10% reduction in the protein content of imported rice (from 2.7g/100g down to 2.4g/100g) would result in an 8% reduction in the protein consumption in Senegal, where the typical protein consumption is only 5% above the World Health Organization (WHO) dietary recommendation. More details are provided in Chapter 2 of this dissertation. Analysis of the data provided by the USDA-FGIS

shows that when the local production has a lower protein content, all exportations are affected the same way regardless of the importer's GPD or buying power. However, countries with a less diverse diet and a small surplus of protein supply might have a part of their population impacted by a deficit of protein supply.

The relationship between protein content and environmental factors such as temperature or water supply has previously been examined at the grain or field scale (Triboi, Martre, & Triboi-Blondel, 2003). The impact of climatic factors at a larger scale has been examined (Rao, Smith, Jandhyala, Papendick, & Par, 1993), and maximum temperature, minimum temperature nine weeks before harvest season has been found to have a significant impact on the protein content of the Soft Winter Wheat (SWW) grown in Idaho, Washington, and Arkansas (Rao, Smith, Jandhyala, Papendick, & Par, 1993).

Obtaining the local and global yield of macronutrients is still a significant challenge. Crop quality is typically monitored by governmental agencies using an in-situ sampling of the crops. The main issue is that there are limited observations on crop quality. For this reason, we argue that nutrition security is still not sufficiently monitored because the change in the quality of the food produced (in terms of calories, proteins, or fat content) is not fully tracked.

Chapter 3 will present a method and dataset for tracking wheat protein content in the United States.

Remote sensing offers a unique opportunity for crop quality monitoring since it has the advantage of being able to provide spatial estimates at multiple times during the growing season. Protein content has been reported to be highly correlated to the ratio vegetation index (RVI) at the grain filling and the anthesis stage of cereals. Preliminary studies by (Wang, Tian, Yao., Zhu, & Cao, 2014) are expected to provide an applicable method for predicting grain yield and grain protein content of wheat and other cereals at a regional scale. Green plants absorb a part of the sunlight through the chlorophyll molecules located in their leaves during the photosynthesis process. Typically, 1% of this solar energy is re-emitted by chlorophylls as fluorescence radiation at 690-800 nm wavelengths. Chlorophyll fluorescence can be used as a proxy to measure plant stress due to environmental constraints. An abnormal metabolism would be translated bu an imbalance between the energy used during the photosynthesis processes and the energy absorbed and re-emitted by the chlorophylls. This fluorescent glow is an excellent indicator of photosynthesis activity and, therefore, the plant's growth. Because this glow is very weak compared to other reflected solar radiations, the usage of satellite-based remote sensing data is relatively new in the quantification of fluorescence. New spectrometers with a very high spectral resolution (0.02-0.05nm full-width half-maximum) centered around 760nm now enable a more accurate measurement of fluorescence at a global level. A study conducted in the US Corn Belt showed a significant correlation between ground-based and satellite measurements (Guanter, et al., 2014). While ground-based measurements exhibit a fine resolution (typically 1 square kilometer), satellite observations are much coarser (typically 3,000 square kilometers), but they were found to be enough to estimate the photosynthetic activity over large crops area. Remote sensing sensors with larger spatial

footprints are suitable in large producers like the United States (444 ha median size) or Canada (1044 ha median size), where crop fields are large and rather uniform so that the spatial resolution of the remote sensing instrument does not have to be as refined (Lowder, Skoet, & Raney, 2016). However, in countries with smaller crop fields such as France (149 ha median size) or China (54 ha median size) remotely sensed data can deliver a mixture of signals from different agricultural landscapes (Meyfroidt, 2017). For these types of regions, finer resolution satellite missions (e.g., Orbiting Carbon Observatory 2 (OCO-2) data) can be utilized as described below. Current and future missions will help in a measure the SIF on smaller or fragmented crops. Missions such as the NASA's Orbiting Carbon Observatory02 (OCO-2) that started in 2014 have a spatial resolution of up to 2km. Current and future missions are summarized in Table 2.

Mission	**Agency**	**Covered period**	**Instrument**	**Resolution (Spatial)**	**Resolution (Temporal)**
FLEX	ESA	2022-	FLORIS	~ 300x300m^2	~ Monthly
Sentinel 5P	ESA	2018-	TROPOMI	7kmx3.5km	1 day
OCO-2	NASA	2014-09-05 Current	OCO-2 OCO SPECTROMETERS	2.25 km x 1.29 km	16 days
GOSAT	JAXA	2014 – Current	TANSO-FTS	1^o x 1^o	Monthly
MetOp-A	ESA	2006 – Current	GOME-2	0.5^o x 0.5^o	Monthly

Table 2 – Remote sensing of Sun-Induced Fluorescence : main satellite missions, time record, and resolution.

These missions will also have a significant contribution to the understanding of the carbon cycle and the improvement of the current models that have been found to have very poor performance when it comes to working on highly managed lands, such as cropland.

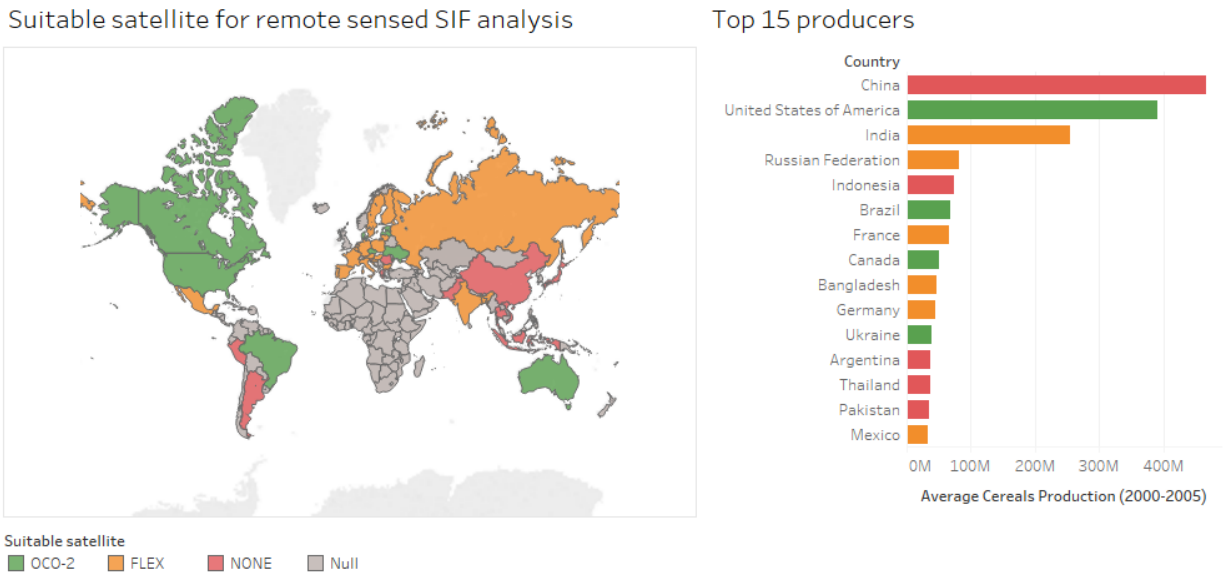

Figure 12 – Suitable satellite mission for using Sun-Induced Fluorescence. The country's colors indicate the current or future satellite missions with a sufficient resolution to analyze most of the cereal crops. We used the median crop size during the 2000-2019 period.

Unfortunately, even the future FLEX missions, purely dedicated to monitoring the SIF, will not have enough spatial resolution to accurately monitor the crops in China, Indonesia, and Argentina, three major cereal producers (Figure 12). We looked at the median crop size in (all croplands) for each country and roughly assumed this median value should be higher than the size of two pixels to get reasonable results. China, the USA, and India are the three main cereal producers (wheat, rice, corn, and soybeans all combined) with very diverse median crop sizes. The USA, major corn, soybean, and wheat producer, has a relatively

large median crop size with several states dedicated to growing one type of cereal (winter wheat for Kansas, corn for Iowa, Soybean for Illinois). India has a median crop size relatively lower than the USA but still has districts dedicated to rice cultivation. China on the other hand, as a major rice and wheat producer, has relatively small crops areas, mainly due to the irrigation requirements of the rice paddies and the farming management. Wheat is moreover often grown coupled with rice and will suffer from the same segmentation of the crops. Therefore, using solely SIF satellite products will not be enough to detect grown areas and differentiate the cereal plants from the surrounding vegetation. Nevertheless, fragmented crops dedicated to rice cultivation in Asia can be easily detected using Sentinel 1A satellite products at 10 meters resolution. Combined with SIF products, there is a great opportunity to detect the growing season and the cultivated area using satellite products.

Chapter 4 is introducing a method to classify rice crops using solely global satellite products and monitor cultivated areas.

Food systems are constantly adjusting, altered by increasing population, shifts in diet, and consolidation of major food producers. Extreme weather events or climatic variations have always impacted food production or even threaten local food security. However, in a globalized economy in which the demand is barely met by the supply, with strong inter-countries and inner-countries disparities, food security cannot be account only for disruption of food production, for instance, due to local weather extremes, but also for nonlocal disruptions, happening on another part of the trade market, but propagating

globally through trade flows. Weather extremes or climatic variation affecting yields, change in policy resulting in a switch in crops produced (for instance from rice to cotton) or the end use of it (for instance corn for bioethanol, corn for feeding), changes in planning and land use resulting in loss of crop area… all these factors can be nonlocal parameters affecting local food security. Furthermore, with rising economies and the near absence of famines within the past decades, food security cannot be reduced to simply dietary energy supply but should also include the quality of the nutrition, for instance, the share of protein in the total dietary energy.

This dissertation is proposing a framework to evaluate the impacts of nonlocal events on local food and nutrition security and consists of four chapters. The first chapter is modeling the propagation and interactions of a major food production deficit with and through the global trade network and quantify the impacts on food supply at the global level. The conversion from food supply to nutrition supply is challenging and requires further refinements: a workflow and dataset created to do so are presented in the second chapter. As lack of available data is a major obstacle, we created models to evaluate the protein content of wheat using machine learning tools and Solar-Induced Chlorophyll Fluorescence data which are presented in the third chapter. As production is the product of yield and harvested area, the fourth chapter is presenting a methodology using remote sensing data, machine learning tools for mapping rice paddies and detecting the change in flooding seasons. The final chapter five is concluding on the work done, the potential application, and future opportunities.

CHAPTER 1

MODELLING THE IMPACTS OF PRODUCTION SHOCKS ON THE GLOBAL TRADE NETWORK AND NUTRITION SUPPLY.

Background & summary

In rainfed agricultural systems, abnormally low rainfall amounts can have dramatic and devastating impacts on farmers' revenues, food prices, and even food availability. Other climatic extremes such as heatwaves, floods also have a dramatic effect on crops productivity (Ding, Hayes, & Widhalm, 2011), (Daryanto, Wang, & Jacinthe, 2015), (Medellín-Azuara, MacEwan, Howitt, Sumner, & Lund, 2016). While the costs of extreme events are typically calculated in the geographical area of direct and physical impacts, the consequences can be felt worldwide. The dependences of some countries' food supply on importations are well-documented. Their vulnerabilities to external food production crises are gaining more and more attention (Lawrence, Lyons, & Wallington, 2010), (Lang & Barling, 2012), (Otero, Pechlaner, & Gurcan, 2013), (Bren d'Amour, Wenz, Kalkuhl, Steckel, & Creutzig, 2016), (Distefano, Laio, Ridolfi, & Shiavo, 2018). Propagations of crises favored by the globalization of commodity trade have been primarily modeled using random graph theory (Bollobas, 2001). This approach is suitable for financial networks (Oatley, Winecoff, Pennock, & Danzman, 2013), (Huang, Vodenska, Havlin, & Stanley, 2013), and can address the risks associated with the spread of a crisis (Battiston, Gatti, Gallegati, & Greenwald, 2012), (Gai & Kapadia, 2010), (Allen & Gale, 2000). Graph theory provides specific measures that are suitable for identifying central players and countries of higher risks. With a global economy increasingly globalized, it gained more attention in recent years for modeling food flows (Baggio, et al., 2016), (Konar, Lin, Ruddell, & Sivapalan, 2018), (Lin, Ruess, Marston, & Konar, 2019). For instance, (Konar, Lin, Ruddell, & Sivapalan, 2018) used network properties, such as connectivity, clustering, and centrality, to better understand both the spatial and temporal transfers of food commodities and vulnerability and

resilience to disturbance. In another study, (Lin, Ruess, Marston, & Konar, 2019) used a similar approach to model food flows within the counties of the USA and highlighted the vulnerabilities of the national food supply chains. Many different indicators have been developed to identify the country at risk of food insecurity following a change in production in a major food-producing country (Jones, et al., 2013). These indicators typically consider a country's dependence on importations, its diversity of providers, its storage capacity, and its buying capacity. However, because food commodities are highly regulated, especially during a food crisis, the complexity and dynamics of international trades cannot be well described by random graphs (Serrano & Boguna, 2003). The general equilibrium theory is also employed to develop agroeconomic models, such as GTAP-W (Berrittella, Hoekstra, Rehdanz, Roson, & Tol, 2007) and IMPACT (Calzadilla, Rehdanz, & Tol, 2010), (IMPACT Development Team, 2012). Partial Equilibrium (PE) or Computable General Equilibrium (CGE) offers the advantage of running short-term or long-term scenarios. These two widely used models are preferred for scenario analysis as they can mimic trade dynamics. However, they suffer from two different challenges. The first challenge being the dependence on elasticities, which might restrict or favor trade, and the second challenge being that these models are mainly static and not dynamic, and, therefore, they can only simulate trade only at one point in time despite that production shocks typically perdure over a couple of years (Nelson, et al., 2013). Another rising challenge in this globalized world is to model the possibility of food substitution, typically represented by a uniform and constant Armington's elasticity coefficient (McDaniel & Balistreri, 2003). CGE also suffers from coarse sectorial resolution (Reilly & Willenbockel, 2010) and large data requirements. However, recent satellite data streams can provide the required data,

and therefore, be combined with the existing theories for modeling propagation of drought impacts through weather, climate, food quality, quantity, and trade information, as we propose here.

Data

Extreme events were characterized using climate data, such as precipitation from the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR), a global precipitation dataset at 0.25-degree resolutions freely available at https://chrsdata.eng.uci.edu/.

Trade information was obtained from the United Nations International Trade Statistics Database (Comtrade), and Food production and consumption were obtained from the Food and Agriculture Organization's datasets (respectively crop dataset and Food Balance Sheets dataset). They are accessible at http://www.fao.org/faostat/en/.

Method

Figure 13 summarizes the empirical model for linking extreme events in one country to nonlocal food and nutrition supply changes in other countries.

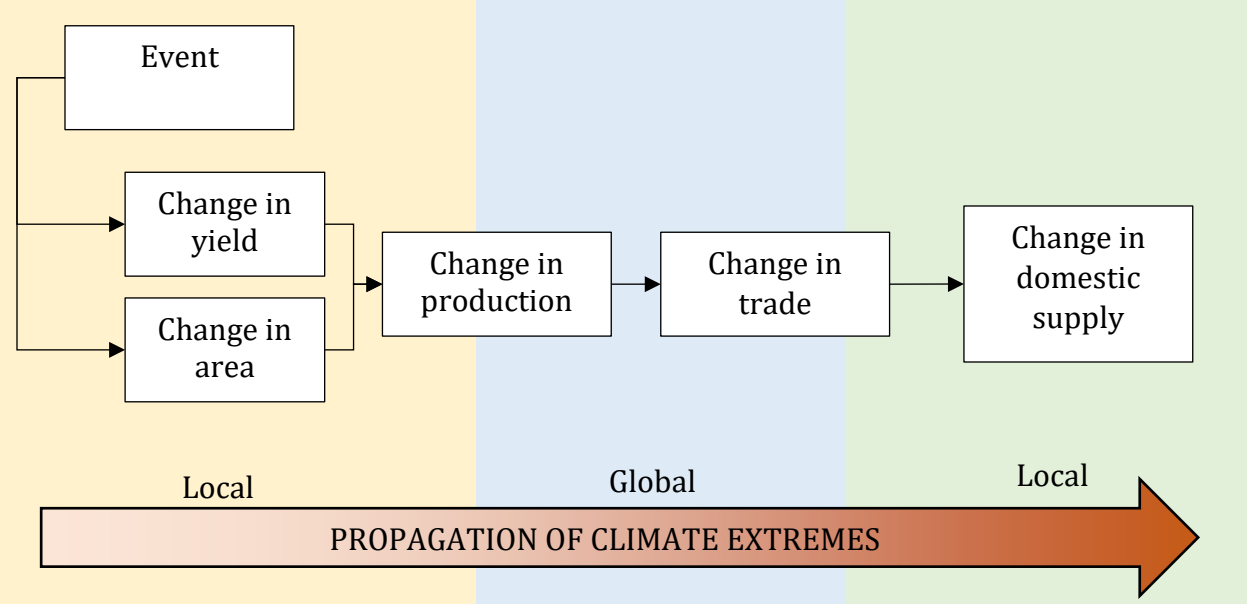

Figure 13 – Proposed empirical model of nonlocal food supply shocks. The background colors refer to the scale considered, i.e., local for the change in food production and food supply, and global for the shift in trade flows.

We characterized droughts using the six months Standard Precipitation Index (SPI) calculated using a nonparametric approach (McKee, Doesken, & Kleist, 1993), (AghaKouchak, et al., 2015). We used the empirical Gringorten plotting position to calculate the marginal probability distribution of the precipitation and standardized them using the normal distribution function such as:

$$SPI = \Phi^{-1}(p_p)$$

$$P_p = \frac{rank - 0.44}{n + 0.12}$$

Where n is the sample size.

While the SPI is an indicator of meteorological drought and other indices specific to agriculture exist, such as the SSI or MSDI (Hao & AghaKouchak, 2014), is chapter aims not to model crop yield. Therefore, using solely the SPI is sufficient to quantify the drought intensity.

Then, using historical data, we calculated anomalies in yields and harvested area using detrended data and converted them into production losses. We propagated these production shocks globally using an empirical network-based trade model calibrated using the past 30 years of trade data. Then, we calculated the changes in food quantity supply globally and deducted the changes in nutrition supply.

We calculated the anomalies in crop yields and areas using the FAO's crop datasets. Trends in yields are typically observed in cereals, mainly thanks to technological progress (farming practice, irrigation, plant species, et cetera). A yield time series may exhibit two or more distinct linear or nonlinear trends due to more than one "revolution" in farming practice or due to changes in data reporting to the FAO. We used least-square spline modeling to fit one or more linear trends and detect potential significant changes in trends over the past 30 years. Figure 14 illustrates such behavior for the corn yield in Brazil and China. For instance, Brazil grows three seasons of maize, and the changes in yield trends are mainly due to the introduction of the second and then third growing season for corn in the 90s and the 2010s (USDA, 2016).

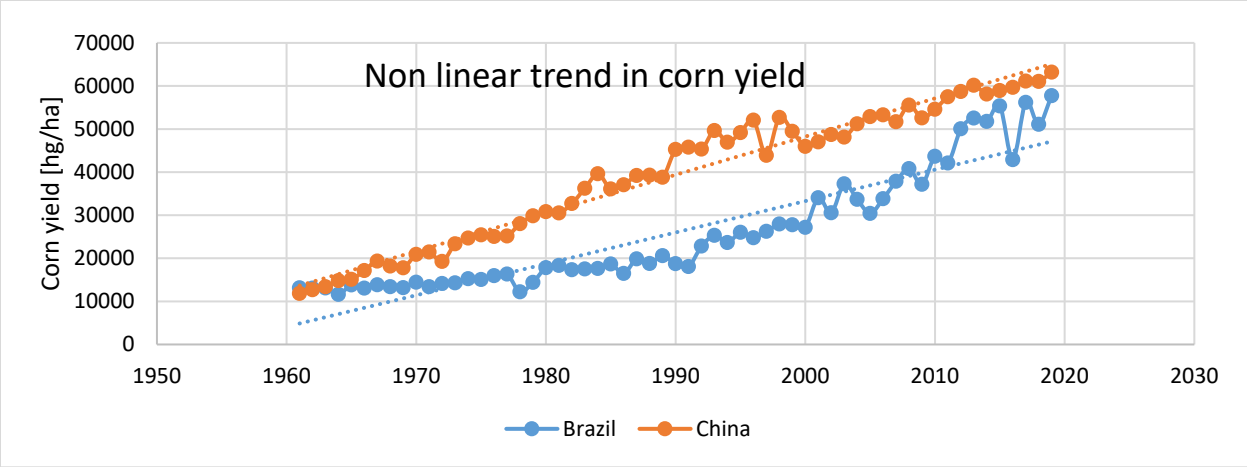

Figure 14 – Example of nonlinear trends in yield : sudden drop followed by a linear trend for China and two linear trends with significantly different slopes for Brazil.

However, our study focuses on major exporters within the past 30 years, and all trends detected in cereal yields were linear. Anomalies were then calculated using the departure from the average of the detrended data, as shown in Figure 15.

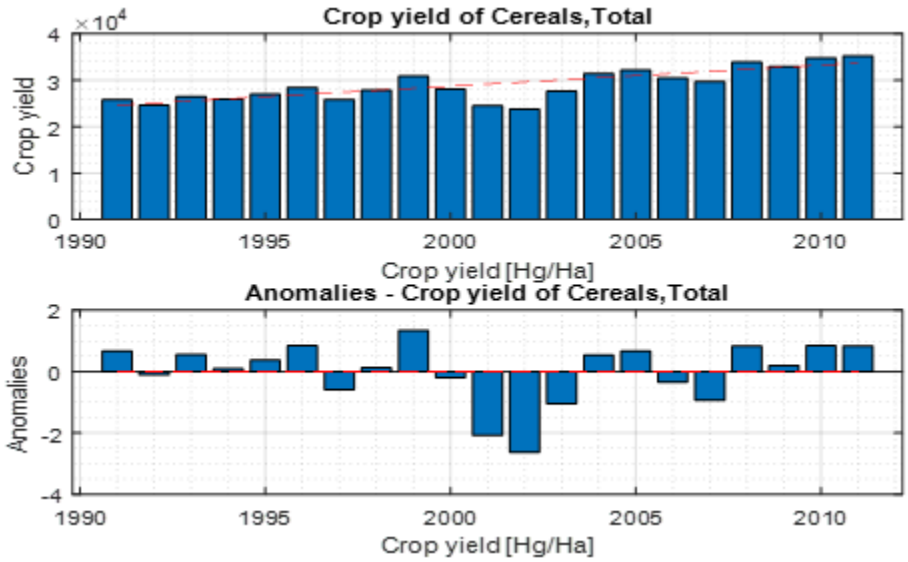

Figure 15 – Quantification of yield anomaly. On the top, the annual crop yield of cereals, the red dash-line represents the trend. On the bottom, departure of the detrended yield from the average value.

The change in total production is then calculated using the yield's anomaly and the harvested area's anomaly as shown in Figure 16 and summarized for each group of commodities in Figure 17. The first figure represents the expected loss of production during the duration of the event.

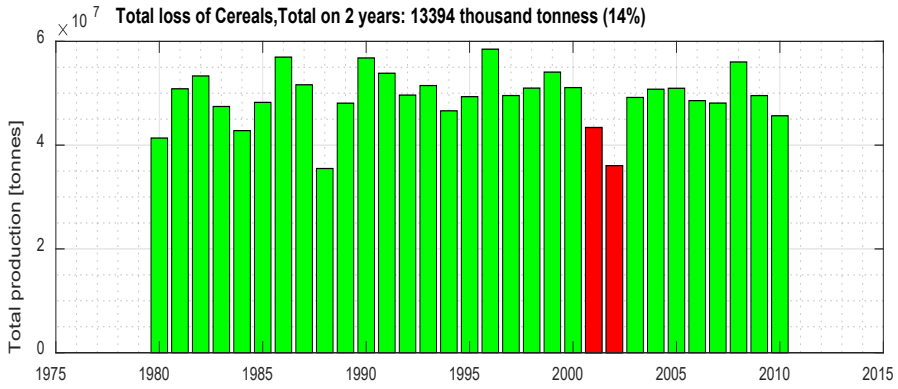

Figure 16 – Example of the expected change in production.

The second figure summarized all the losses in production for each commodity gathered into a food group. The change in food quantity expressed in tonnes is also converted into a change of nutrition supply expressed in tonnes of proteins and calories.

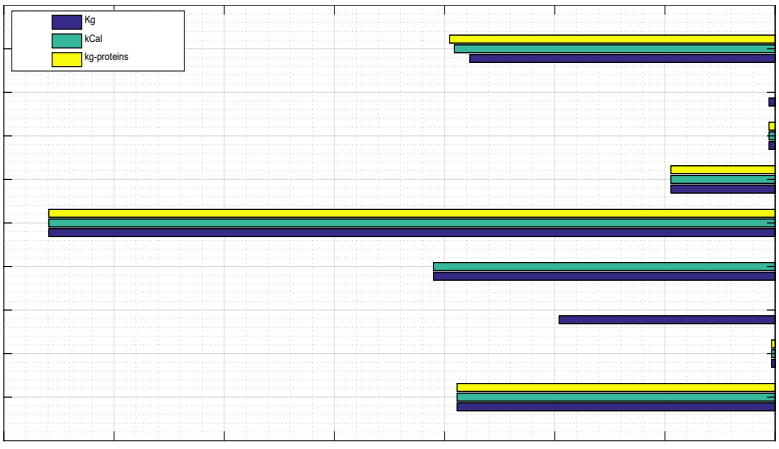

Figure 17 – Example of the total change in production for all food groups.

These changes in production are translated into changes in exportations. Figure 18 shows the production and trade deficit (normalized by the maximum production deficit) of corn products from Canada.

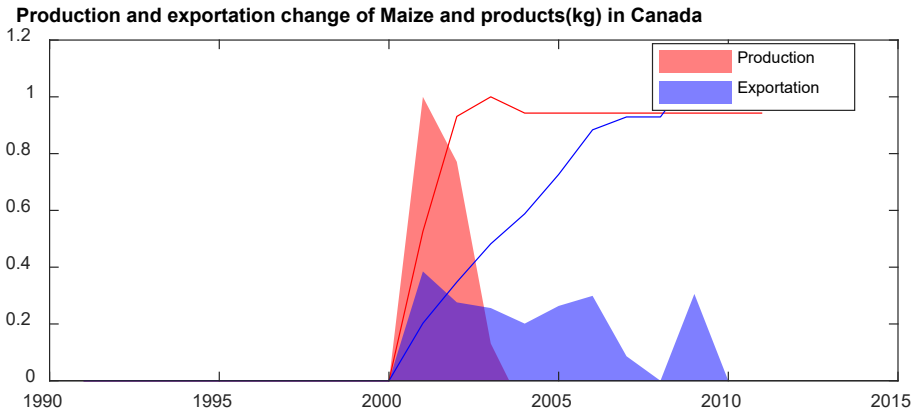

Figure 18 – Visualization of the export propensity coefficient. The blue and red areas indicated the normalized annual decreases in exportation and production of corn. The lines represent the cumulative changes.

It reveals that the 2001-2003 deficit in production was 80% absorbed by using the storage, which was rebuilt slowly until 2006. The ratio of cumulative production deficit by the cumulative trade deficit is used as a proxy to estimate how a country has been using its storage capacity during the event period. Figure 19 indicates the actual changes in exportations. The next section explains how historical data were used to calculate the export propensity coefficient and to be able to convert a change of production into a change of exportation.

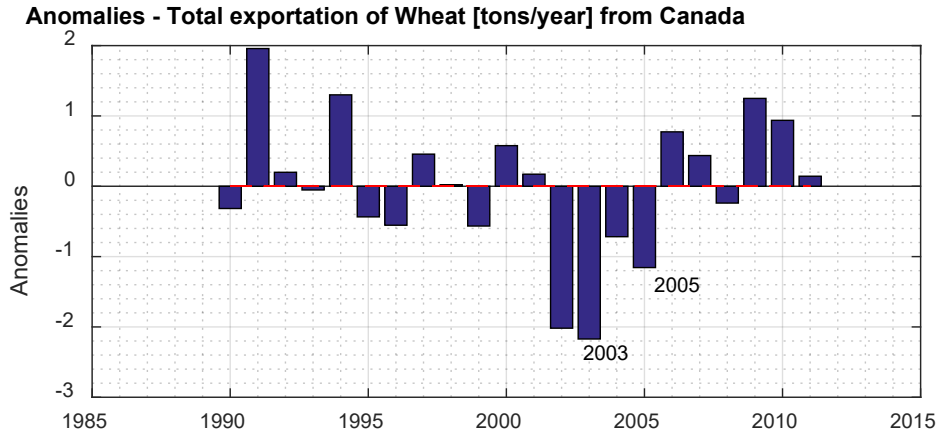

Figure 19 – Anomalies in exportation of wheat from Canada. The anomalies were standardized.

Production shocks are then propagated through an empirical network model. The model is propagating a country's production shock to its storage capacities and exportation capacities. One production shock for one country is becoming an importation shock to its commercial partners, who will propagate it the same way until convergence as illustrated in Figure 20.

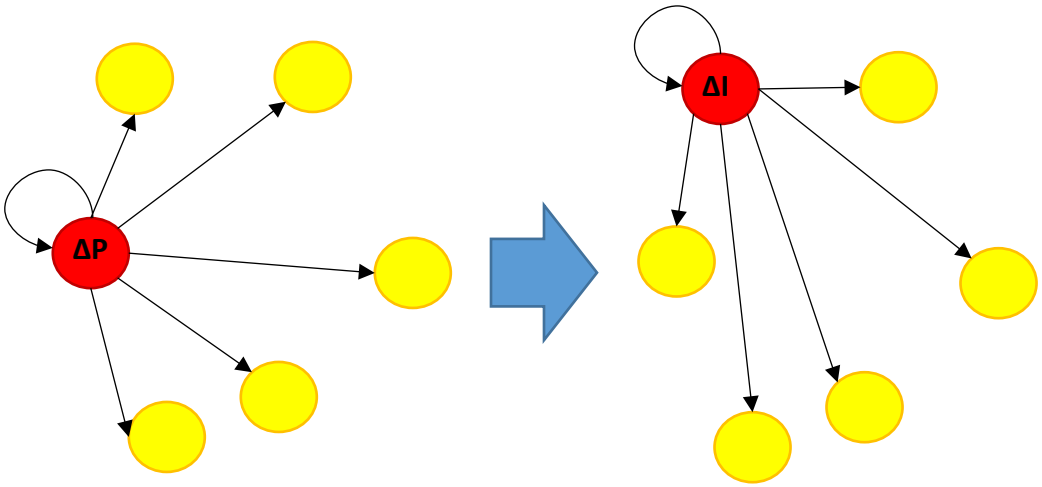

Figure 20 – Proposed empirical network-based trade model. In red, the country experiencing a shock (production or importation) and in yellow, its commercial partners.

We also consider all commodities independent of each other's, which means we are not modeling replacement of one commodity by another one in importers' portfolio. We tested the hypothesis of eating more meat and fewer cereals during food crises but haven't found a significant correlation between the two. Figure 21 however displays such behavior for two small insular countries (the Fijis and Vanuatu), which exhibit both a strong correlation between corn supply and bovine meat supply within the past 50 years. This correlation is however in the longer term and there is no indication that it is the result of a sudden food crisis.

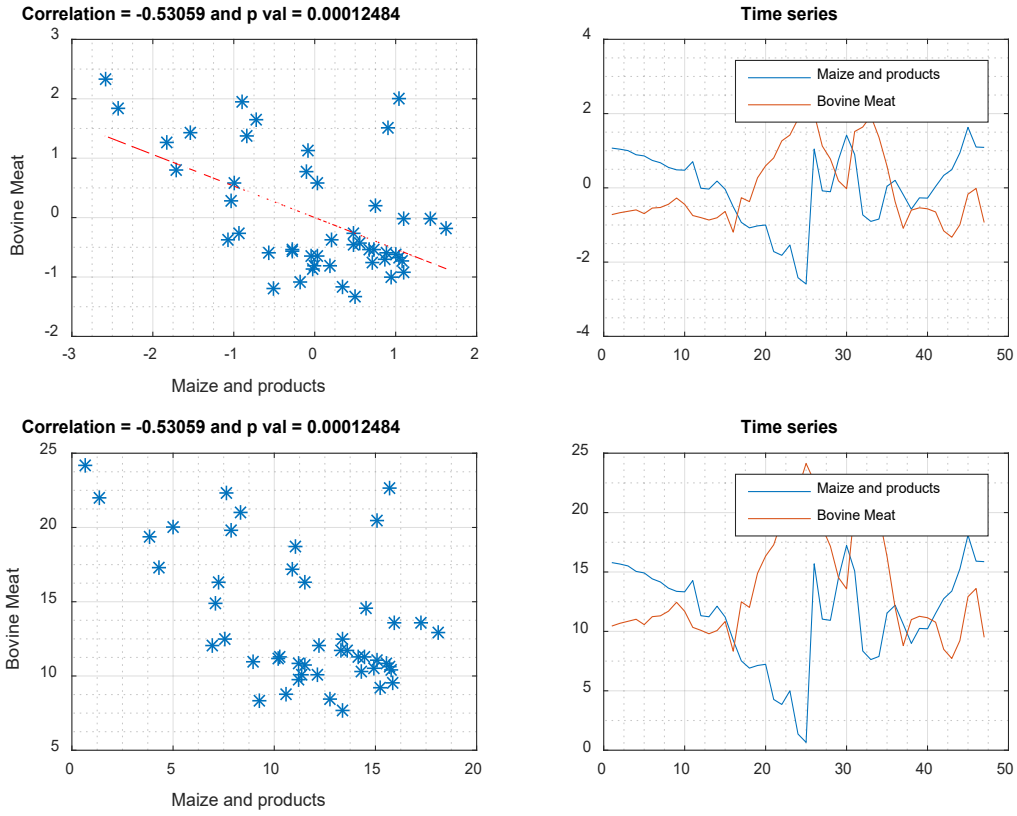

Figure 21 – Replacement of maize and bovine meat in the domestic supply

For each commodity c, we define a production loss in the country k during the year τ_0:

$$\Delta_c^k = P_C^k(\tau_0) - P^{*k}_C$$

$$\Delta_c^k = \beta \cdot P_c^k(\tau_0)$$

Where P^* is the expected production and is calculated using the area harvested and the deviation of the crop yield from its long-term detrended mean.

$$P_c^{*k} = A \cdot \left(\langle Y_c^k \rangle - Y_c^k(\tau_0)\right)$$

We assume that the change in production will impact both the domestic supply and the exportations. We assume that there will be no change in importation. Analysis of major production shocks from the major producers did not exhibit significant changes in importations but rather changes in storage utilization and exportation of the surplus. The expected value of total exports and accumulation can thus be calculated using:

$$E_c^k = E_c^k(\tau_0) - \alpha^k \cdot \Delta_c^k$$

and

$$A_c^k = A_c^k(\tau_0) - (1 - \alpha^k) \cdot \Delta_c^k$$

The coefficient α^k is the export propensity of the country k and ranges from 0 to 1. A coefficient of 1 means that all the production shock has been translated into an equal decrease of exportation while a coefficient of 0 means that the production shock has been reflected into the domestic supply. It is calculated as:

$$\alpha^k = \frac{E_c^k(\tau_0)}{E_c^k(\tau_0) + A_c^k(\tau_0)}$$

The spillover of the production shock to the exportation of the element c from the country k to the country l is calculated as:

$$T_c^{k,l}(t) = T_c^{k,l}(\tau_0) - \alpha^k \cdot \Delta_c^k \cdot \frac{T_c^{k,l}(\tau_0)}{E_c^k(\tau_0)}$$

Each country l will now have a deficit of importation Δ_c^l, where

$$\Delta_c^l = T_c^{k,l} - T_c^{k,l}(\tau_0)$$

And this deficit of importation will act and propagate the same way as the initial production shock. Therefore, we have an iterative process to calculate the propagation of every change in the trade until the solution converges to the final value.

$$\left| \frac{T_c^{k,l}(t_n)}{T_c^{k,l}(t_{n-1})} \right| < 1 - \epsilon$$

Convergence is double-checked using a global balance of the commodity c such as:

$$P_c^k(t_F) - P_c^k(\tau_0) = \sum_l [A_c^k(t_F) - A_c^k(\tau_0)]$$

The export propensity coefficient can be seen as the storage policy and is limited by potential export bans.

Results

We analyzed several historical events, such as the 2002 North American drought, the 2003 French drought/heatwave, and the 2009 Argentinian drought, as summarized in Table 3.

Table 3 – Summary of the events analyzed.

Event name	**Intensity**	**Production shock**	**Trade shock**	**Most impacted countries**
2002 North American drought	SPI: -1.5	13%	60%	Middle East Central America
2003 French drought / heat wave	SPI: -2.0	14%	60%	West Africa Southern Europe
2009 Argentinian drought	SPI: -1.3	4%	40%	South America

We identified the 2002 North American drought as a severe extreme event affecting food production in several states, including Nebraska, Iowa, Colorado, Wyoming, and Utah in the United States and Alberta, Manitoba, and Saskatchewan in Canada. Figure 22 reveals the extent and intensity of the event, and Figure 23 reveals the final impacts on nutrition supply globally.

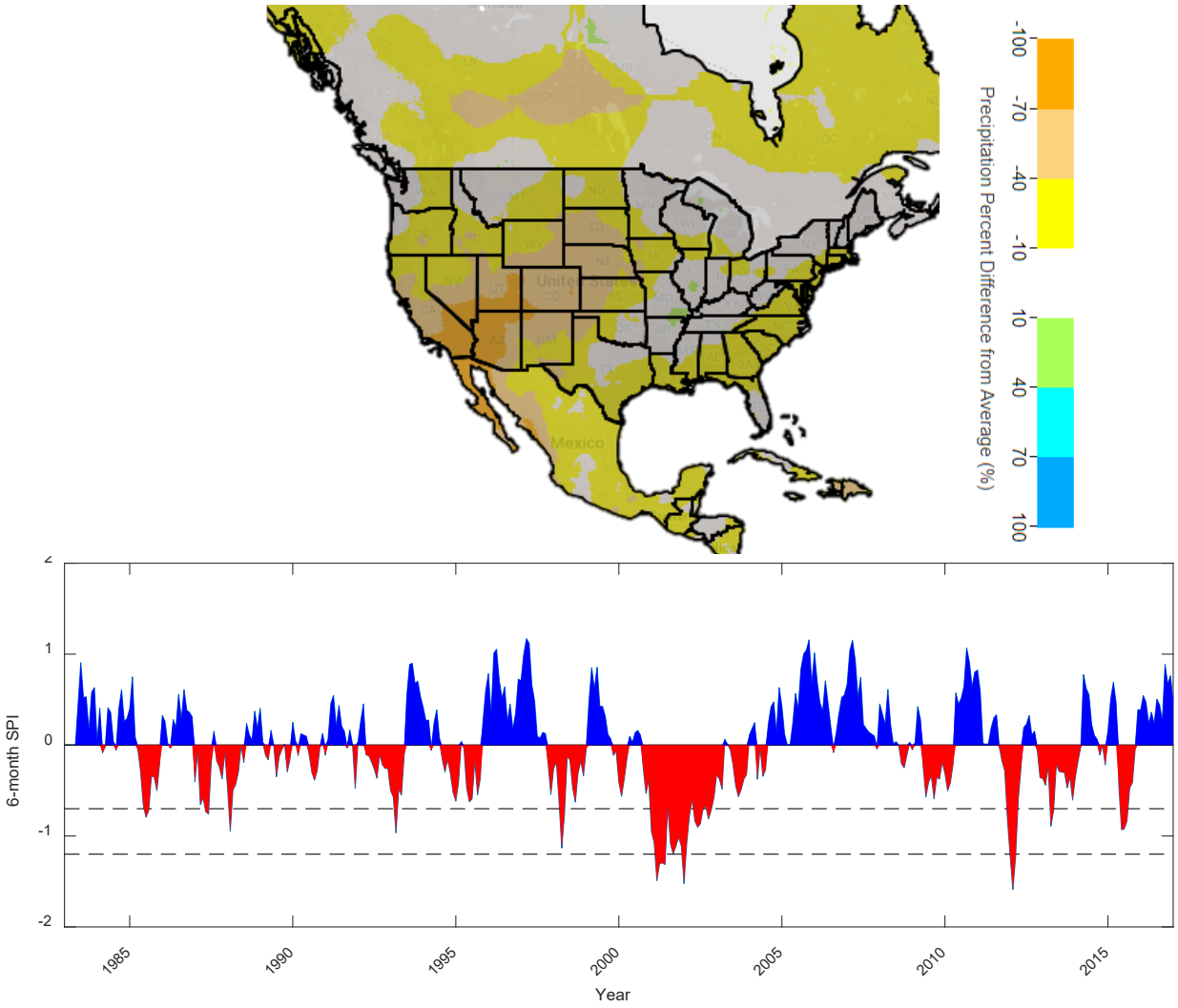

Figure 22 – Intensity of the 2002 North American drought. The top plot represents the extent of the drought using anomalies in Spring precipitation using the TerraClimate precipitation product. The bottom part represents the area-weighted SPI time series for the region.

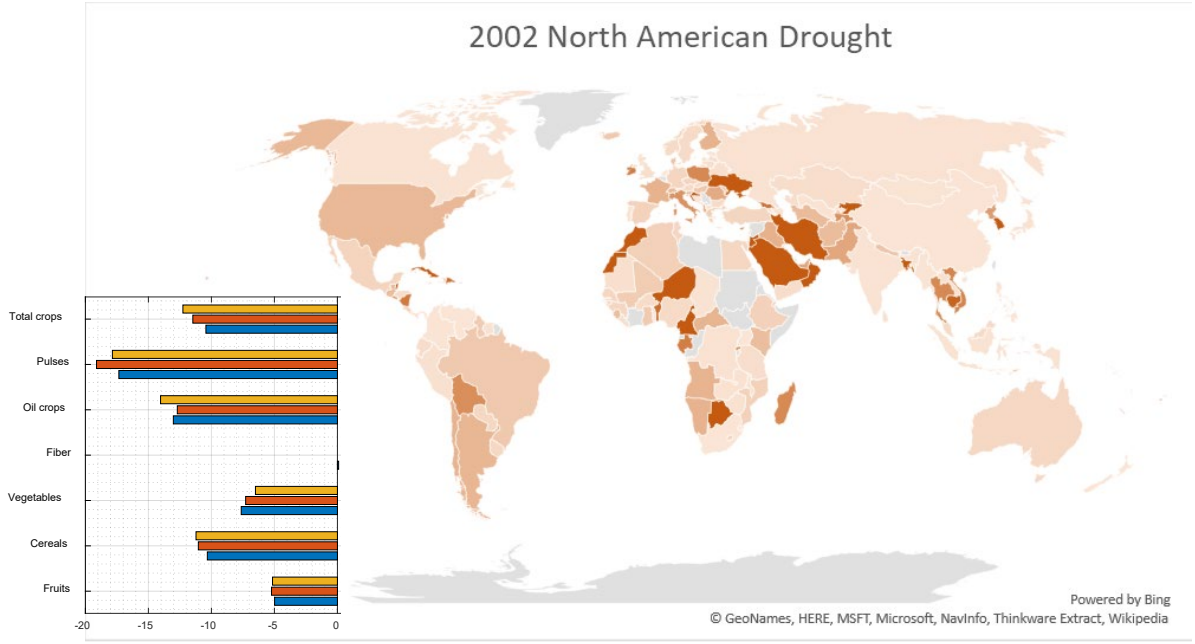

Figure 23 – Change in nutrition supply due to the 2002 North American drought. The bar plots represent the percentage change in production (total quantity in yellow, calories equivalent in orange, and protein equivalent in blue).

Pulses, oil crops, and cereals were the most affected food group in the USA and Canada, with up to20% of loss of production for the pulse group. Overall, the event resulted in a loss of 12.5% of total mass, or 11% and 10% calories and protein equivalent. The countries the most affected are in the Middle East region (Saudi Arabia, Iran, Oman, Afghanistan), Asia (Cambogia, Thailand), Central America (Guatemala, Honduras), South America (Bolivia, Guyana, Venezuela), African (Niger, Morocco). Interestingly, the United States is also impacted (by a decrease of Canadian exports).

Figure 24 displays the extent and intensity of the drought in France during the 2003 West European drought/heatwave. This event and its impacts on primary productivity and food production is well documented in the literature (De Bono, Peduzzi, Kluser, & Giuliani, 2004), (Ciais, et al., 2005), (van der Velde, Wriedt, & Bouraoui, 2010) and Figure 25 reveals the impacts on the global nutrition supply.

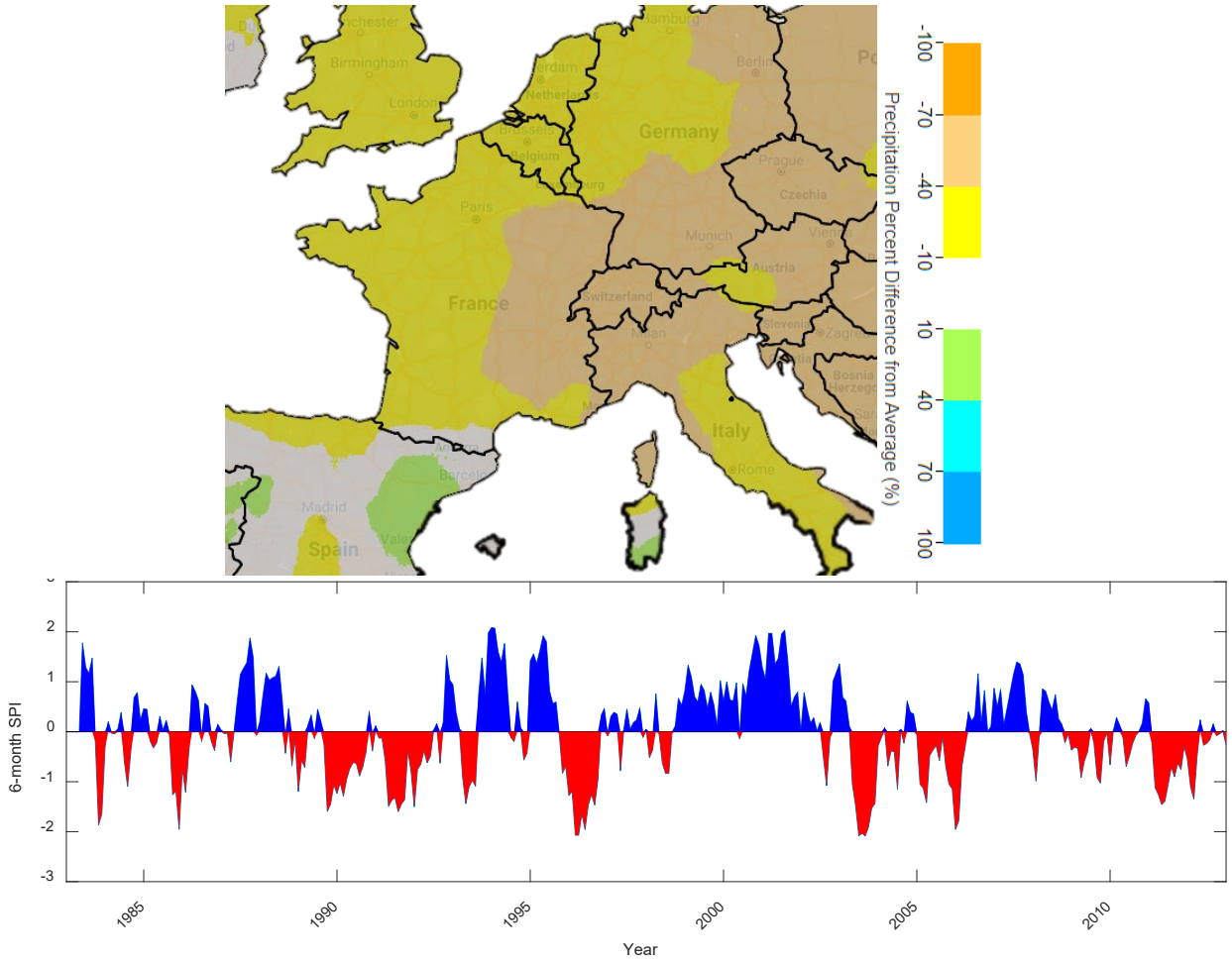

Figure 24 – Intensity of the 2002 North American drought. The top plot represents the extent of the drought using anomalies in Spring precipitation using the TerraClimate precipitation product. The bottom part represents the area-weighted SPI time series for the region.

Pulse, cereals, and fruits groups are the most impacted, with up to 15% loss of production for the cereal groups. Overall, roughly 10-12% of food and equivalent calories and protein were lost due to the drought/heatwave. The countries the most impacted were Northern Africa (Algeria, Morocco), West Africa (Senegal, Mali), Europe (Spain, the Netherlands, Italy).

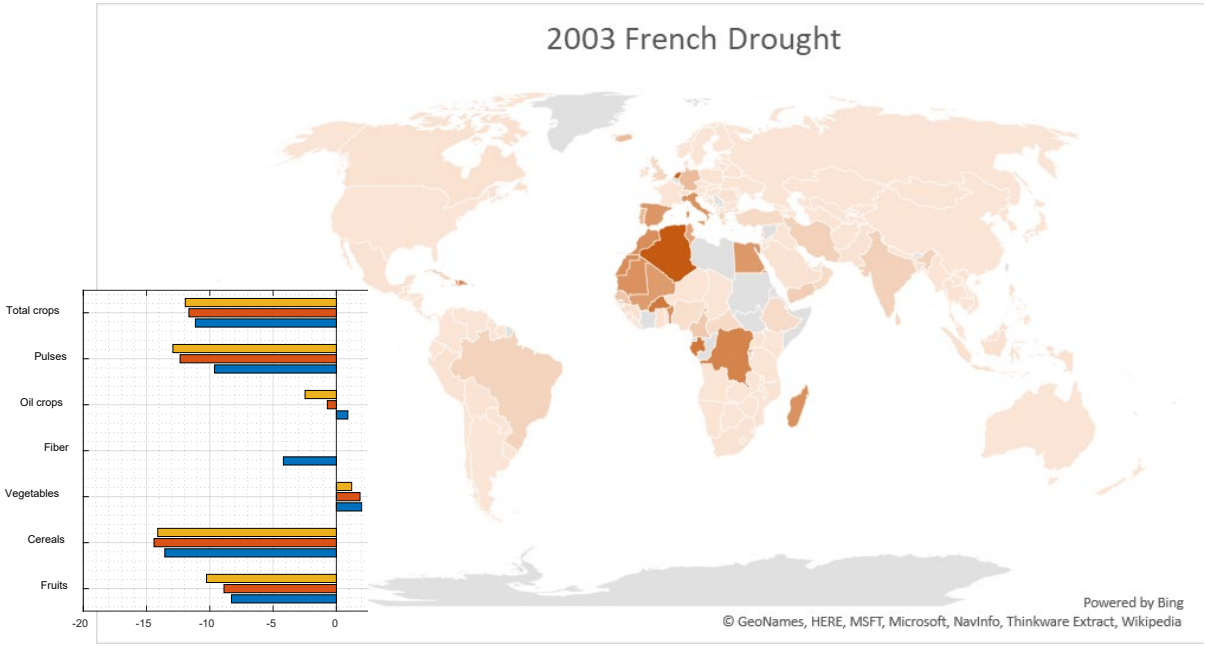

Figure 25 – Change in nutrition supply due to the 2003 French drought. The map represents the change in calories supply due to the 2003 French drought. The bars plot on the left represents the change of production of several food groups.

Figure 26 displays the extent of the drought in most of the Southern American region, including Argentina. Corns and soybeans were reported to be the most affected crops (Sgroi, Lovino, Berbery, & Müller, 2021). Cornfields were more vulnerable to pests due to the water stress, and the flowering of soybeans dramatically decreased due to the

combined water and extreme heat stress (USDA Foreign Agricultural Service Office of Global Analysis, 2009). Figure 27 reveals the impacts on nutrition supply globally.

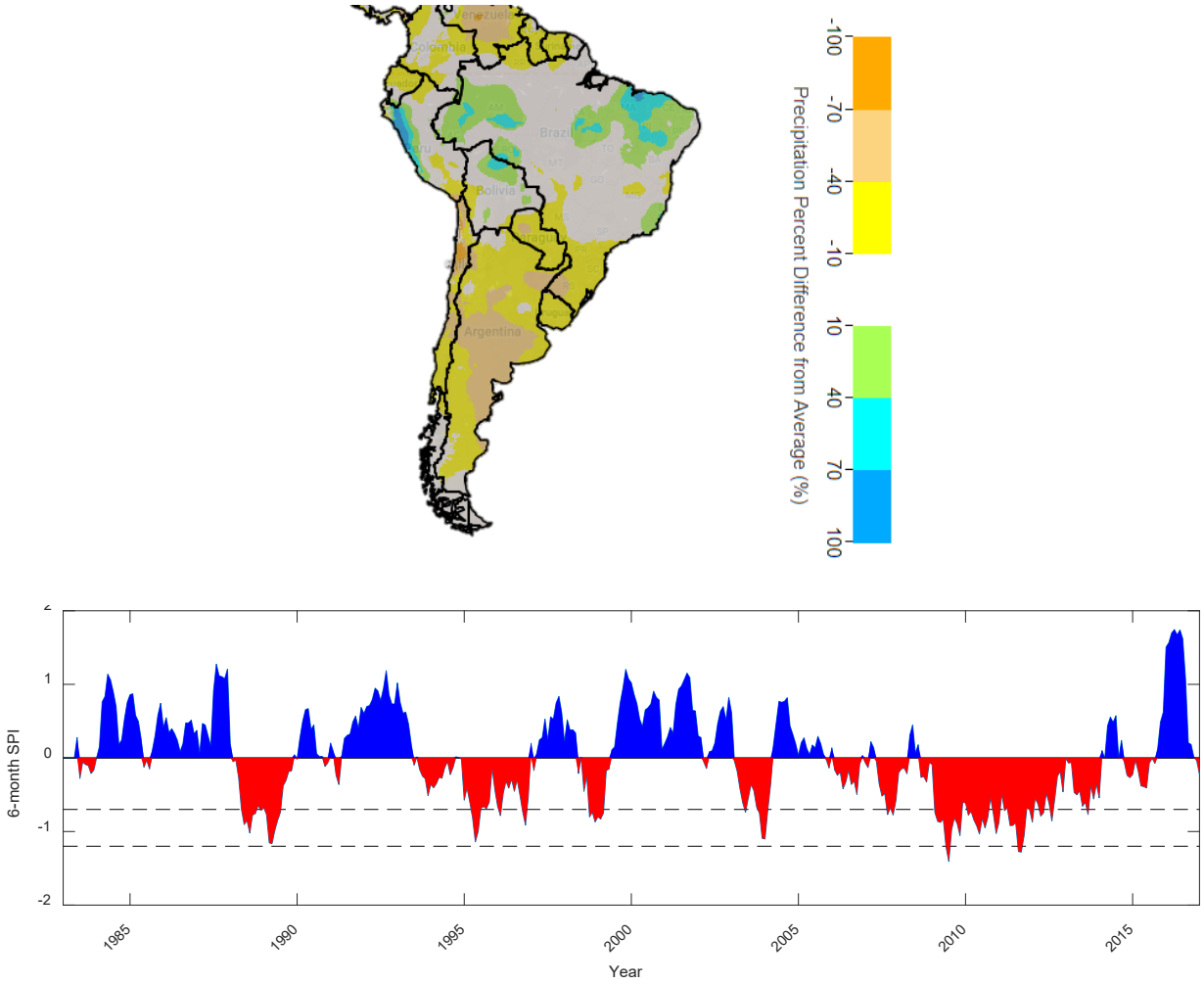

Figure 26 – Intensity of the 2009 Argentinian drought. The top plot represents the extent of the drought using anomalies in Spring precipitation using the TerraClimate precipitation product. The bottom part represents the area-weighted SPI time series for the region.

Oil crops, which include soybeans, were among the most impacted food groups, with about a 7% loss of production and equivalent calories and protein produced. The overall event resulted in about a 4% decrease in calories and protein produced. The countries the most

impacted are neighbor countries in South America and Germany in Europe, the whole Middle East, and South Africa.

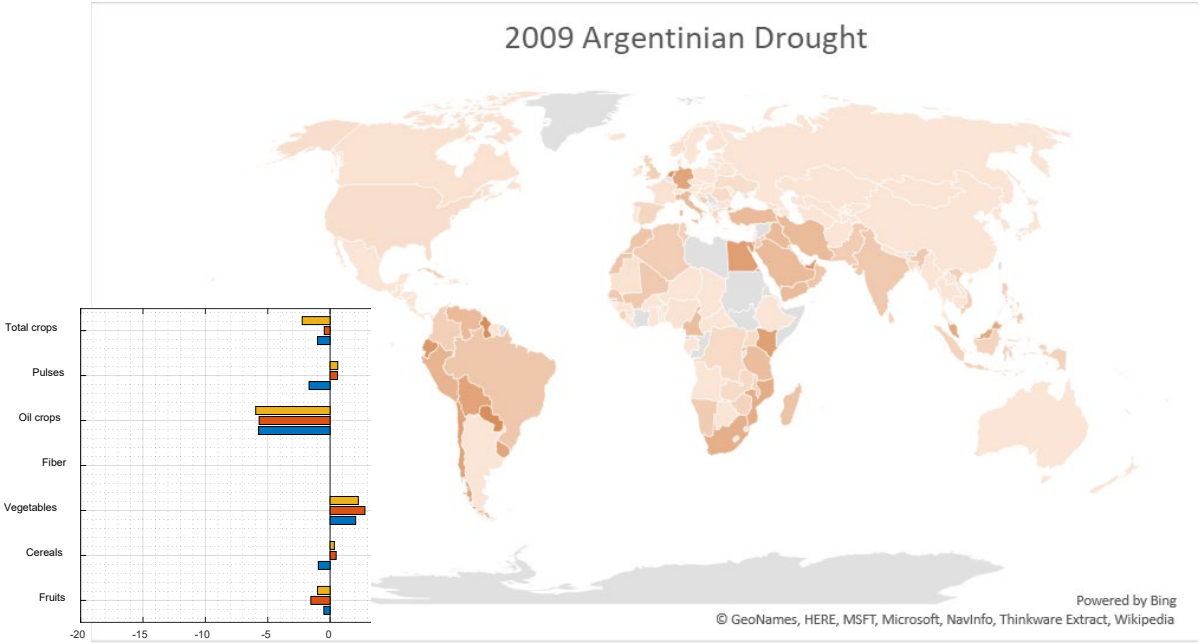

Figure 27 – Change in nutrition supply due to the 2009 Argentinian drought. The map represents the change in calories supply due to the 2009 Argentinian drought. The bars plot on the left represents the change of production of several food groups.

Discussion

We compared our findings with the best information available in the scientific literature or governmental organization's reports available to us. The results of the Argentinian drought can seem counterintuitive at first, as the loss of production calculated in our analysis is much lower than expected for a drought of such intensity. However, it was reported that farmers used the stunted corn crop for feeding the livestock (NASA Earth Observatory, 2009), and the corn was harvested as livestock forage rather than for the grains. Furthermore, soybeans are mainly used for the feeding industry or oil for

bioenergy, which explains the globally low impact on food and nutrition supply. It was also reported that the severe heat killed 1.5 million head of cattle (Sgroi, Lovino, Berbery, & Müller, 2021) which consequently disturbed the meat market as Argentina is a major meat producer and exporter. The loss of production in France is corroborated by several reports (COPA-COGECA, 2003), (UNEP, 2004) as well as the loss of production in North America (Svoboda, et al., 2002), (Agricultural and Agri-Food Canada, 2002). The results are also consistent with the trade flows analysis shown in Figure 3 and the supplier diversity analysis shown in Figure 4 in the introduction and reveal two types of impacts: first order impacts, due to the direct decrease of exportation from one major producer to its preferred commercial partners (e.g., North America to the Middle East), and second order impacts due to the disruption of the global trade network (e.g. North America to West and South Africa).

However, our empirical model has several limits to consider when doing scenario analysis. For instance, we do not consider export bans that one country could impose due to a disruption in the global trade network to protect its domestic supply. Such measures are popular to stabilize domestic prices, and 33 countries used 87 of these bans during the 2007/2008 and 2010/11 food price crises (FAO, 2011). Apart from the Indian ban on non-basmati rice in 2007 and the Russian ban on wheat exports in 2011, such bans from large exporters are relatively rare, and our model is already capturing the decreases of exportations due to market dynamics through the export propensity coefficient. Moreover, we did not consider price modeling, a non-negligible component, as a food crisis could occur from pure speculation. For instance, speculation is thought to have partially caused the 2007-2008 food crisis (Ghosh, 2009). While there are models able to take price

dynamics into account – IMPACT model from the IFPRI (Rosegrant, et al., 2008), the focus of this study is not the interactions between food prices and food supply but rather the consequences of a nonlocal production shock in food unsecured countries.

Furthermore, extreme weather events are often reduced to their main impact, i.e., negative impacts on the crop yield. In the next chapters, we demonstrate that extreme events can also affect the harvested areas (Chapter 4) as well as the nutritional values of the cereals (Chapter 3), and we demonstrate that such variation can also have significant impacts on nutrition security (Chapter 2).

Conclusion

In an increasingly globalized world, domestics food systems should not be considered individually but globally as a whole network. Such globalization and intensification of trade flows have supported countries to resist to local food crisis but exposed them to a nonlocal production crisis. In this study, we modeled the propagation of production shocks from large exporters through the international trade network to its commercial partners. We identified countries directly impacted by the decrease of exportation as well as countries indirectly impacted by the disruption of the global trade network. We identified the countries the most vulnerable to the nonlocal food crisis.

CHAPTER 2

A NEW MACRO-NUTRITION SUPPLY DATASET (MANU).

Background & Summary

Identifying long-term trends and temporary changes in human nutrition is a challenge due to limited access to reliable and annual data. Induvial or household surveys, which are the most complete and detailed, are usually conducted once or twice per decade only and on a small population, not always representative of a country., and are commonly lacking in low-income countries. At the country level, national statistics are less informative as they do not provide information on the disparities within the country but are usually available each year for most countries. The Food and Agriculture Organization (FAO)'s Food Balance Sheet is a dataset of food available for human consumption, i.e., the residual of a food balance equation using relatively reliable data on production, storage, trade, processing, seeding, and feeding, and estimate on food waste. Because of its long-term coverage since 1961, the FAO's Food Balance Sheet is one of the most widely used datasets of food and nutrition available for human consumption. However, it presents a few drawbacks, such as losing information on the origin of the food supplied (local production or importation), and the macronutrient supply estimated relied on conversion factors which are assumed uniform and invariant.

Analyzing trends and change in global nutrition is becoming more and more important in a globalized food system, where cheap food can be easily produced, exported, and can shift the diet of other countries (Drewnowski & Popkin, 1997). Converting food quantity to nutrition quantity is not straightforward. In the current FAOSTAT models, the first limitation is that the final product, Food Balance Sheets, does not include information related to the geographical origin of the food available in each country. The second

limitation is that the conversion factors used to convert food quantity to nutrition quantity are assumed to be constant and uniform despite evidence of substantial spatial and temporal variations due to environmental, cultural, and climate conditions. Wheat, one of the most consumed commodities, has, for instance, a protein content ranging from 8g/100g to 16g/100g (median value at country scale), (Smika & Greb, 1973), (Terman, 1979), (Casagrande, David, Valantin-Morison, Makowski, & Jeuffroy, 2009). The main reasons for such large variabilities are climate-related (extreme events such as drought and heat wave), environmental constraints (soil quality and pollutants), or culture's methods (e.g., fertilizers, crop calendar, cultivar used). Such variability in protein content due to extreme climate conditions might overestimate the protein supply in countries at-risk and lead to aggravated undernutrition. Similar variation can be found for rice (Cruz, Cagampang, & Juliano, 1970), (Li, et al., 2018).

This study's objective is to offer a unique dataset for studying variability in nutrition supply worldwide. We developed a dataset of the food supply with crucial features, such as protein, fat, carbohydrate content, geographic origin, and year of production of the food. The latter allows tracking how nutrition moves around the world from major producers to consumers. The study is limited to macro-nutrients (protein, fat, carbohydrate) and does not consider micro-nutrients since the latter are extremely sensitive to the utilization of the food (storage, cooking style, et cetera). We also estimated the confidence intervals of the nutrition supplied for every country and evaluated them against the MDER (Minimum Dietary Energy Requirement), ADER (Average Dietary Energy Requirement), and DES (Dietary Energy Supply). The algorithm, data, and toolbox can be used to calculate the nutrition supply in a country with more accuracy and calculate the uncertainty associated

with these estimates. We define calory-secured as the dietary energy supply (DES in kCal/cap/day) being higher than the Minimum Dietary Energy Requirement (MDER in kCal/cap/day)). The MDER being an index calculated every five years by the WHO, we average the DES over the same period as the MDER.

Knowing the ranges of nutritional values for all commodities is nearly impossible. In this study, we limited our efforts to the most produced, traded, and consumed commodities (i.e., cereals). Table 4 summarizes the ranges of values documented in the literature. The number between crochet [ ] indicates the standard value used in most studies and provided in the FAOSTAT's dataset. Carbohydrate content is usually not directly available and calculated using the energy balance:

$$Calories^{Total} = Calories^{Protein} + Calories^{Fat} + Calories^{Carbohydrate}$$

Protein, carbohydrates, and fat can be easily converted to energy knowing that one gram of proteins or carbohydrates provides four calories, and one gram of fat provides nine calories. The ranges of protein, fat, and carbohydrate have been selected from the literature, and the range of calories has been calculated using the variations of one parameter only. For instance, for wheat, the protein content can vary from 8 to 16 g/100g resulting in a range of calories from 316 to 348 Cal/100g, the carb content can vary from 60 to 75g/100g resulting in a range of calories from 310 to 370 Cal/100g, and finally, the range of fat can vary from 1.1 to 3.2 g/100g resulting in a range of calories from 323 to 342 Cal/100g. Therefore, the range of calories was assumed to be from 310 to 370 Cal/100g due to variation in carbohydrate content only. The change in the kernel composition may affect all three parameters simultaneously, but we assume independence due to a lack of

data on the relationships between the three parameters. Annual data availability is one major challenge, and currently, only a few countries offer yearly harvest quality reports with such data. Therefore, we are using an only range of values in this study. More information on data availability can be found in Appendix 2.2.

Table 4 – Ranges of nutritional values of the most consumed cereals.

Commodity	**Nutritional value**	**Reference**
Winter wheat [100g]	Protein: 8 - 16 [12.5%] Carb: 60 - 75 [66%] Fat: 1.1 - 3.2 [2.3%] Calories:310-370 [334 Cal]	Protein: Chapter 3 Carb: (Steenfeldt, 2001), (Pirgozliev, et al., 2003) Fat: (Dror, Rimon, & Vaida, 2020) Calories: Calculated
Corn [100g]	Protein: 6.8 - 12.5 [9.5%] Carb: 78 - 85 [70%] Fat: 4 - 6 [4.3%] Calories: 346-374 [356 Cal]	Protein: (US Grains Council, 2021) Carb: (Hopkins, 1899) Fat: (Hopkins, 1899) Calories: Calculated
Rice (Milled) [100g]	Protein: 6 - 13 [6.7%] Carb: 74 – 80 [81.7%] Fat: 0.6 – 2.9 [0.7%] Calories: 357 – 385 [360 Cal]	Protein: (Hamaker, 1994) Carb: (Puri, Dhillon, & Sodhi, 2014) Fat: (Puri, Dhillon, & Sodhi, 2014) Calories: Calculated

We consider low and high protein in the following sections, respectively, as 10.5 g-protein/100g and 14.5 g-protein/100g rather than the standard 12.5 g-protein/100g for wheat.

Methods

We used the production, trade, and food balance datasets from the FAOSTAT database. The domestic food supply for human consumption is calculated using the residual of a mass balance for each of the 426 commodities.

$$Supply = Production + Import + \Delta Stock - Export - (Feed + Seed) - Waste$$

Each commodity has a nutritional value defined by a standard number, which is uniform and stationary, or an average number or a distribution that is specific to a year and a country. Figure 28 shows illustrate the algorithm used to create the dataset.

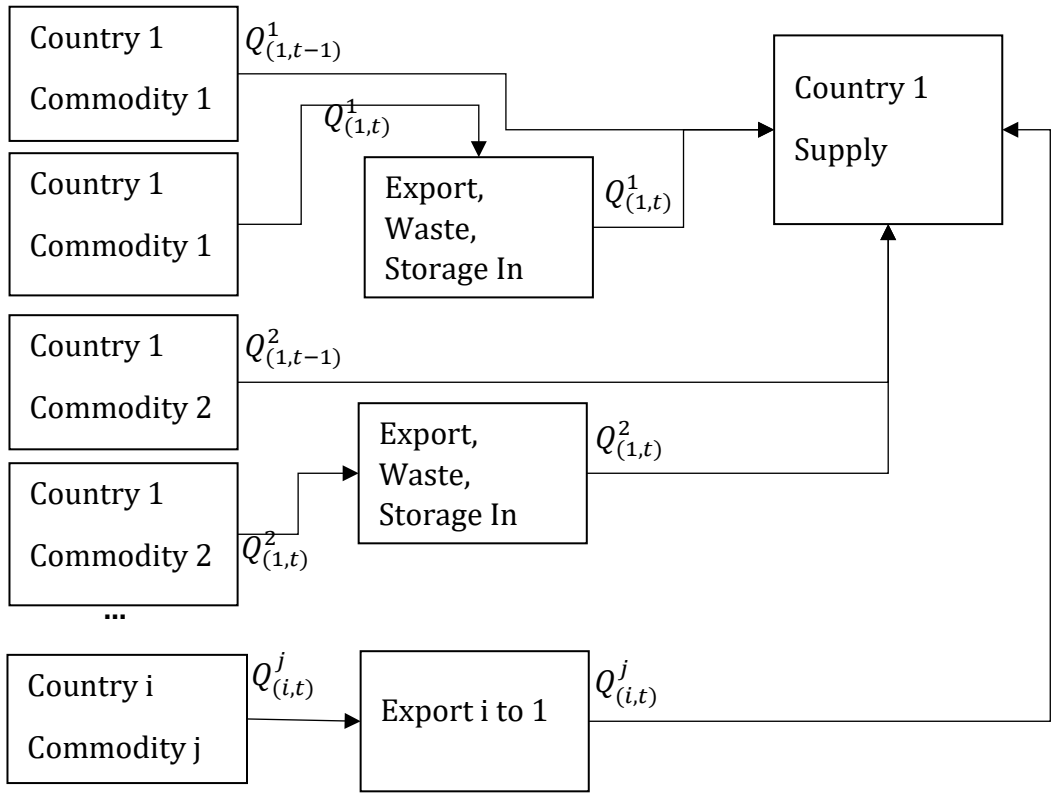

Figure 28 – Nutrition supply's dataset workflow. Q_(i,t)^j represents the nutritional value of the commodity j produced in the country i during the year t. It can be a number estimated, surveyed, or a distribution of values.

Results

Using the methodology previously described, we calculated the nutrition supply dataset for each country, commodity, and year available. Using the uncertainties ranges from Table 4, we evaluated the uncertainties in nutrition supply by varying one or several food items and changing the values in one or several countries.

For instance, Figure 29 exposes the impacts of 10% change in the US Wheat protein content on calory security in 2010-2014. Yellow color represents countries switching from calory-unsecured to secured when the US Wheat's protein content is increased by 15%. Orange color represents countries switching from calory-secured to unsecured when the US Wheat's protein content is decreased by 15%. Green and red countries are calory-secured and respectively unsecured regardless of the US Wheat scenario. The figure on the right reveals the most impacted countries.

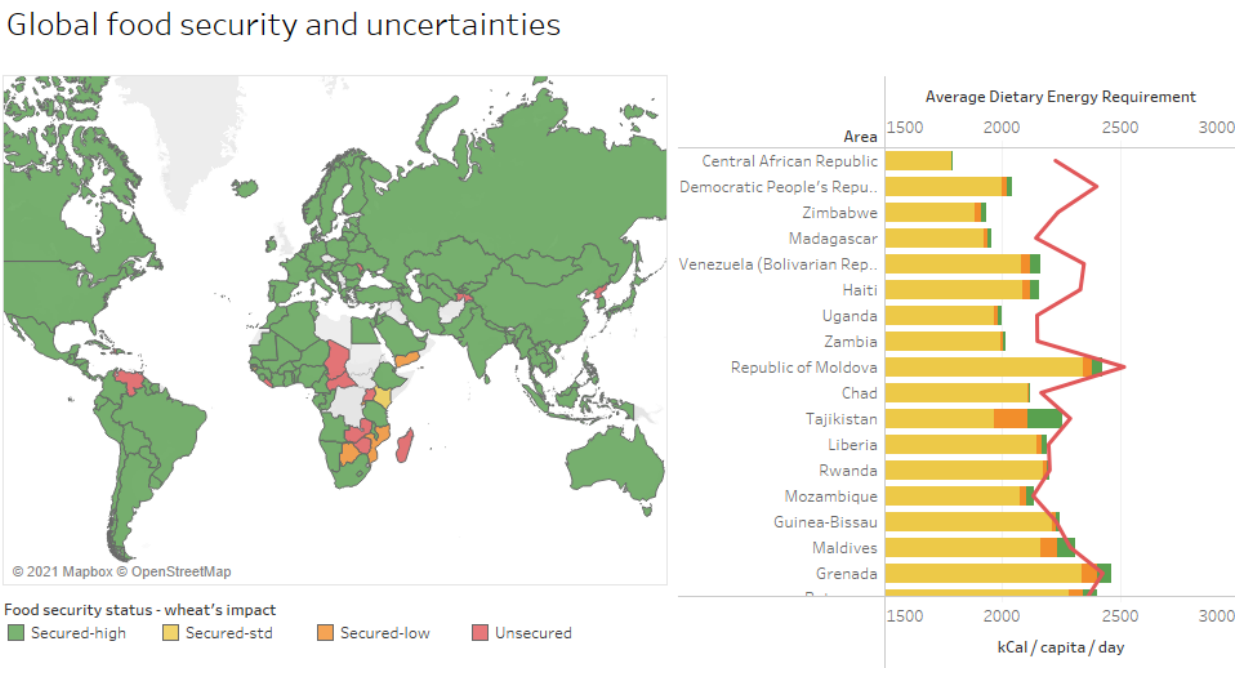

Figure 29 – Significance of variations in US wheat's quality to the protein supply and impact of a change of 15% in wheat protein content on the global dietary energy supply (DES). Left: Global maps of food security: In red, the country's DES is smaller than the MDER. In Orange, the country's DES is smaller than the MDER if the US Wheat has 10% less protein than the standard value, in yellow the country's DES is higher than the MDER if the US Wheat has 10% more protein than its standard value. Right: The top food-unsecured countries with the MDER (red line) and the DES (Yellow).

Figure 30 reveals for each country the impact of a change in wheat quality imported on their protein supply security. Here we show the impact of one grade lower (10.1 g-protein/100g instead of 12.5g-protein/100g) wheat protein content on the corresponding protein supply. The results show that most of the impacts are felt in Central and Latin America, which heavily rely on cereal imports from the United States, and from Western Africa and the Middle East, which rely heavily respectively on European and North American productions. The change in protein supply could be as high as 34% for Belize

going from 72g/cap/day of protein supplied using standard values down to 46g/dap/day using a 10.5g-protein/100g-US-wheat.

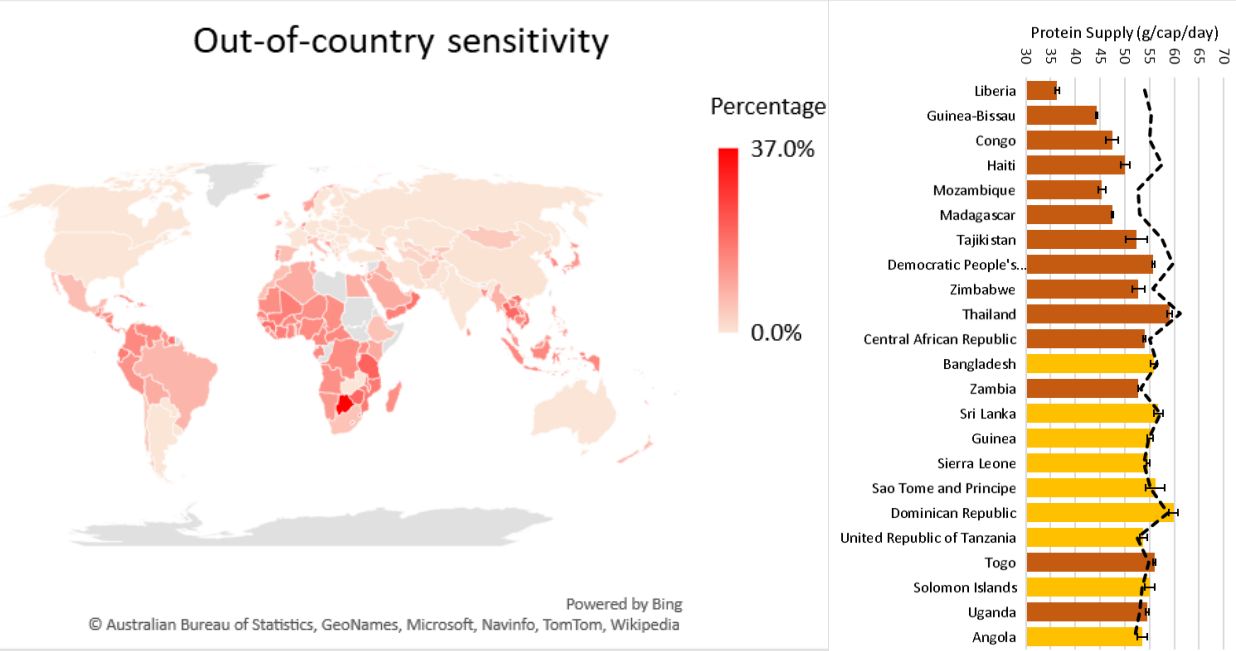

Figure 30 – Significance of variations in nonlocal wheat's quality to the protein supply: impact of a change of 15% of the wheat protein content in traded wheat products on the global protein supply security. Left: Percentage changes of protein supply when out-of-country wheat loses 15% protein content. Right: Countries ranked by their deficit of protein supply. In yellow, the deficit or excess is within the uncertainty created by out-of-country imports of wheat.

Discussion

The results suggest that variations in the nutritional composition of the food can result in important changes in dietary energy supply. When using the national average of dietary energy supply, one can consider one country as meeting the minimum nutrition supply requirements using standard nutritional values but not meeting these requirements when using the lower quality products. We found for instance that departure of protein of

cereals from their standard values can induce important variation in the estimation of energy and protein supply. By looking at one of the most traded commodities, changing the wheat protein content of the US Wheat only will have a strong impact on the calories supply in other countries (using 4 calories per gram of protein). Looking at an out-of-state variation, i.e., for each country, we lower the wheat protein content imported only, reveals that several countries (e.g., Thailand, Republican Central Africa, and Zambia) could have a protein supply over the dietary requirements if all their wheat exports were 15% higher protein content than the standard values. In an opposite direction, several countries (e.g., Belize, Zambia) have a protein supply over the recommendations assuming a standard value of wheat, but below the recommendations when assuming the uncertainties in all imports. Most other countries are either not affected at all, or their dietary recommendations are within the uncertainty level of their protein supply.

In this study, we looked at the national average (at country scale) without considering the socio-economic aspects and disparities inside a country. A large part of a population can be hungry or undernourished even if the domestic dietary supply is higher than the domestic dietary recommendations due to disparities in buying power, access to food, et cetera. Moreover, food can be processed and lose some of its nutritional value. This is one potential source of uncertainty in our study however Figure 59 in appendix 2.1 is showing the food utilization by country for the main cereals. It demonstrates that most of the food consumed in the countries mentioned in these studies are mainly destined to direct food for human consumption, therefore we are minimizing the errors.

One key challenge is to allocate flows of food to storage, seeds, consumption, … Storage in a given year might include food produced in the same year or a previous year or imported. Here, we assumed that the storage is fully emptied each year and distributed to seed, feed, export, waste, food, and refilled gain by the production and importation. We also assumed that the food nutritional profile of each commodity will remain the same during all steps: storage, trade, waste, feed/seed, and supply. We note that storage conditions can impact the nutritional quality of the commodities and it is also conceivable that high-grade commodities are stored more often rather than low-grade ones. However, given the lack of detailed data, we assume the same nutritional profile for produced and stored food.

Trade can also be a factor to consider as different countries have different usages for their imports and different standards. We decided however to assume the same nutritional profile for all shipments, regardless of the country of destination, resulting in the same profile for domestic usage and all exports. Our analysis of the export profiles for US Wheat obtained from USDA-FGIS is supporting this assumption. Figure 31 reveals the average protein content of the wheat exported by the United States in 2019. The data are provided by the USDA Federal Grain Inspection Services (FGIS) on roughly 200,000 shipments to 106 countries from 2010 to 2020. The shipments will be used for domestic supply in other countries, i.e., direct human consumption, processing, feeding, seeding, or will be wasted.

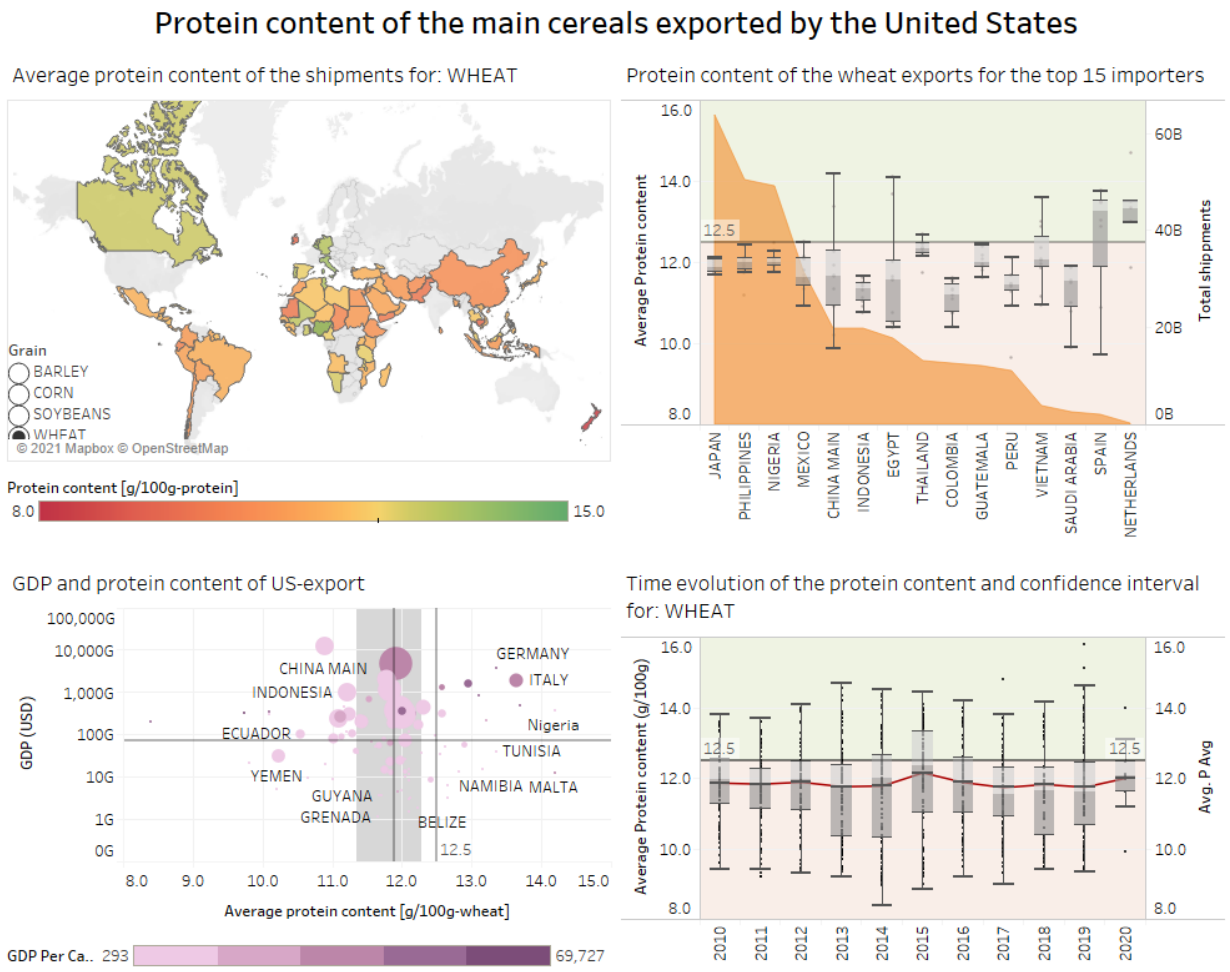

Figure 31 – Overview of US wheat export: quality, quantity, and destination. Protein content of the US Wheat Exportations in 2019, data from USDA FGIS. The top left map represents the countries and the average protein content of the wheat exports, the two right charts display the average protein content of all exports by country and by year, and the bottom-left chart classifies countries by GDP. In all charts, the 12.5% protein content lines represent the standard value of wheat protein content.

Although there can be large variations of protein content between years or countries (from 8 to 14 g/100g-wheat), the variations are randomly distributed as shown on the bottom left scatter plots. The y-axis is the GDP, the color indicates the Per Capita

GDP, and the size of the circle represents the total volume imported. The scatter plot is randomly distributed suggesting that there is no relationship between the GDP and the protein content of the exports. Nevertheless, the median protein content of the wheat exported is almost systematically under the 12.5g/100g-wheat widely used in the literature, which is why we decided to consider the nutritional values of the food consumed in addition to the quantity of the food consumed for our food security analysis. The top left map shows the average protein content exported to each country (average 2010 to 2019). Among the richest nations, Canada and the European Union have consistently received shipments with a higher value of protein content, more suitable for bread making, while New Zealand is received wheat with much lower protein content, probably destined for feeding their livestock. The top-right chart represents the range of protein content of the wheat for the top 15 importers in volume. China, Mexico, Egypt, and Spain are among the countries with the largest variation in protein content. The bottom right chart reveals the same variation of protein content but for each year. Although the median value is relatively stable, the 95% confidence interval can vary a lot, especially from 2013 to 2016 following the 2012 US drought.

Final usage is also an important factor to consider. Lower-grade cereals are for instance usually used for feeding animals, while the highest grades are used for human consumptions. Here we assumed no relationship between the nutritional value and the final usage. Finally, food waste, as defined by the FAO, occurs at all steps from production to consumption. For instance, wastes can happen during the production stage, even if the meteorological conditions were ideal. During the storage time, waste will most likely occur on the lower quality grade which has a shorten lifetime although it is possible that only the

highest-grade commodities are stored and that the lowest one is consumed immediately. The same can happen for the traded communities (import-export operations). We however assumed here that the wastes commodities have the same grade distribution as the produced commodities.

Conclusion

In this chapter, we created a dataset keeping key information on the food consumed at the country level: the geographical origin and year of production. This way, we can link the food consumed to a place of production, and, in case of lower nutritional values, we can assess the impacts of one event affecting food quality in the United States to nutrition security in Central Republic Africa or Belize. We showed that our dataset can better estimate nutrition security at a country level and that current nutrition and food security studies using standard values of protein content for cereals might overestimate food security in many regions of the world. This type of information is critical as it can be used to assess a country's risk of undernutrition due to extreme events outside of its borders. It can also help countries to improve their food security status by buying higher protein content food. In our analysis, we have demonstrated that there were no significant differences of quality between the harvest and export of wheat from the United States and Canada. However, further analysis incorporating commodities price, socioeconomic data at country scale need to be conducted to further refine nutrition supply estimate and the inner variations, especially for the poorest or most vulnerable populations inside a country. One key limit to such analysis is the data availability. Currently, only a few countries offer public harvest quality reports each year, and often time such reports are not in a user-

friendly format easy to analyze. Updating a database of wheat quality would then be time-consuming and fastidious. However, there are strong opportunities to use remotely sensed data to predict the harvest quality eat year when harvest reports are not available to easily accessible. This is the object of the next chapter.

CHAPTER 3

PREDICTING PROTEIN CONTENT IN HARD RED WINTER WHEAT.

Background and summary

Cereals are an important part of human nutrition and are widely used as a source of protein (Prosekov & Ivanova, 2018), (Belton & Taylor, 2004). In some countries, wheat and derived products are the main sources of protein supply as demonstrated in Chapter 2. Protein contents are expressed in crude protein concentration based on a 12% moisture basis for grain in the United States (where a value of 12% is reasonable) and based on the dry matter basis in Europe (which corresponds to 13.5%-DMB). It is not measured directly but rather derived from the concentration of nitrogen in the grain using a conversion factor. Being able to measure this value is important to farmers, to help them determine their revenue and support their marketing decisions (Moschini & Hennessy, 2001), to dealers and to mills that require a constant quality of wheat, and to consumers. To date, there is no global dataset publicly available of wheat protein content, other than annual reports for grains corporations or inspection reports of exportations from public authorities in several countries.

In this chapter, we are proposing an algorithm to build a dataset of annual wheat protein content for several countries, and we are evaluating several data-driven models to predict wheat protein content using globally available remote sensing data.

In the United States, wheat is classified into six classes based on the hardness and the growing calendar. The Hard Red Winter wheat (HRW), which has a high protein content, is mainly grown in the Great Plaints. The Hard Red Spring (HRS) wheat and Durum (D) wheat are grown in the Midwest, the Soft Red Winter (SRW) wheat is grown in the Southeast, the Soft Winter (SW) wheat is grown in the Pacific Northwest regions, and the

hard white (HDWH) wheat is grown on the West Coast (U.S. Department of Agriculture, 1957). The Hard Red Winter (HRW) is the most cultivated in the United States (about half of the production) as seen in Figure 32 and in Kansas (about 95% of the production). Furthermore, HRW and HRS represent respectively approximately 50% and 40% of the total US Wheat Exportation.

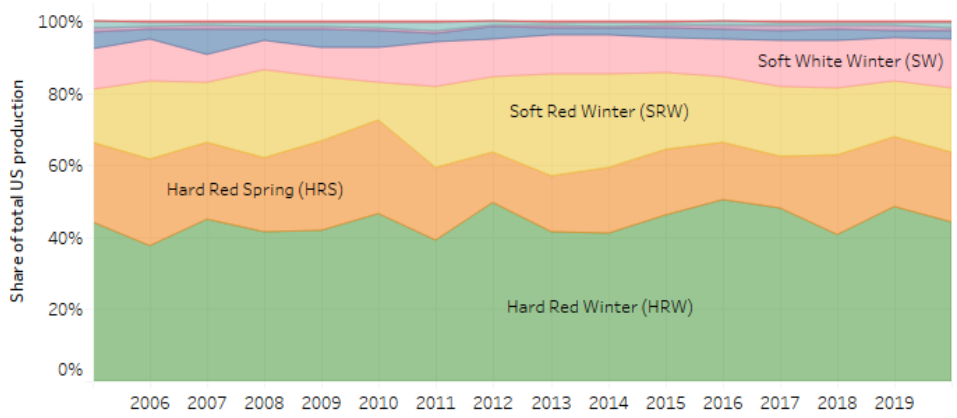

Figure 32 – Wheat production in the US by wheat class (2005-2020).

HRW typically has a high protein content and presents excellent milling and baking characteristics. Protein content varies depending on the plantar used, the growing conditions (soil properties, fertilizer), and the climatic conditions (precipitation, temperature) (Moschini & Hennessy, 2001), (Porter & Semenov, 2005), (Jarvis, et al., 2008). Figure 33 reveals the variations in protein content from one year to another and within States in the United States for HRW wheat. For instance, in 2015, the protein content was relatively low but also relatively uniform was in 2013 and 2014, both values and spread were higher. The green and red background colors represent the values

respectively over or under the 12.5% protein content, which is a widely used standard value.

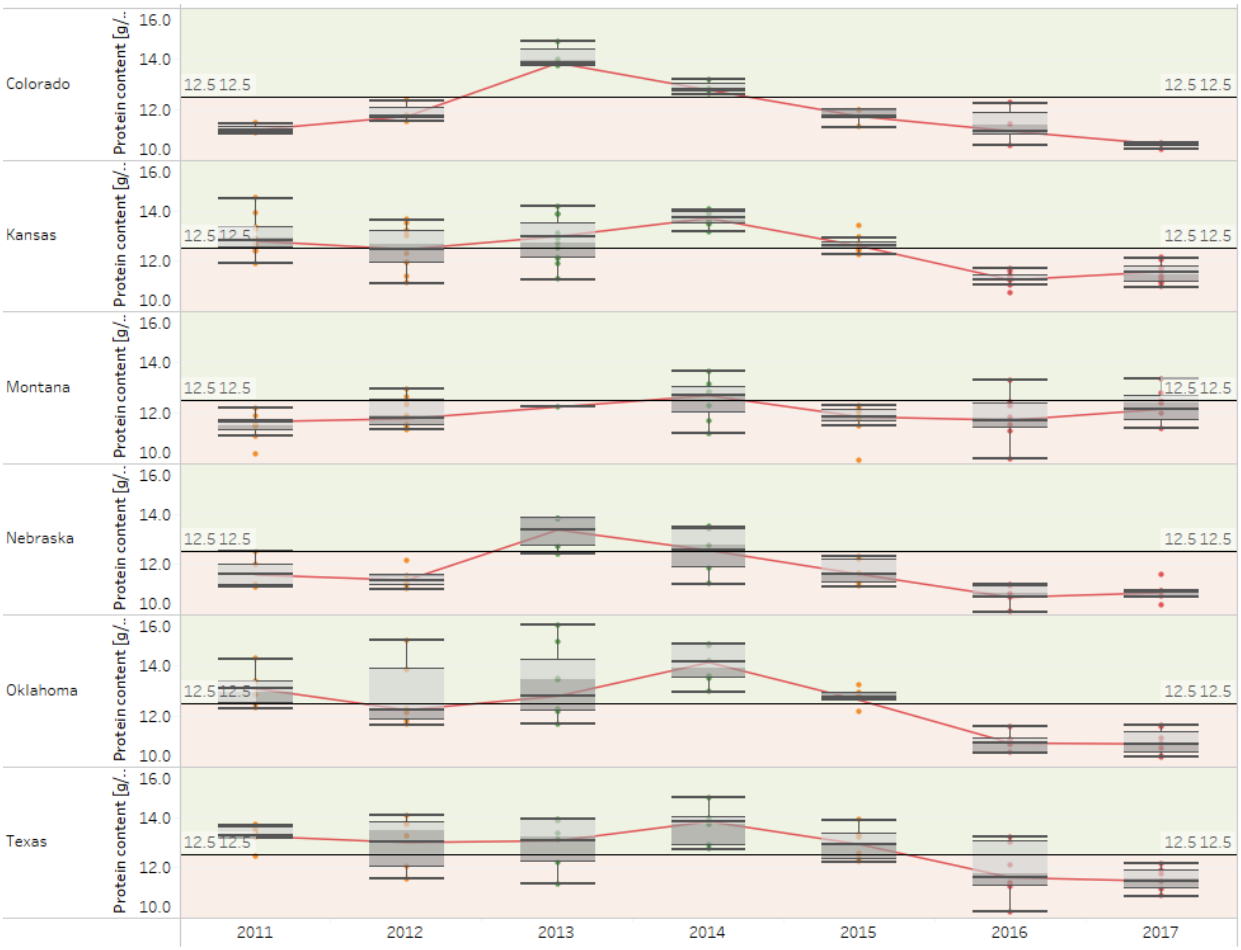

Figure 33 – US Wheat's protein content by state and year. Box plot of the wheat protein content in the top US producing states from 2011 to 2017. The 12.5% horizontal lines represent the standard wheat protein content value.

Understanding the physical processes relevant to wheat protein content is fundamental to developing a data-driven prediction model. Grain storage proteins (GSPs) are mostly synthesized from amino acids, i.e., grain nitrogen, which comes mostly from the remobilization of the plant nitrogen to the grain (roughly 60% to 95%) and from the post-

anthesis uptake of nitrogen. Grain storage protein is therefore closely related to the N-uptake in the stem and leaves pre-anthesis and to the N-transport post-anthesis from the vegetative organs of the plant to the grain (Dupont & Altenbach, 2003), (Kichey, Hirel, Heumez, Dubois, & Le Gouis, 2007). After flowering, it is suggested that the N uptake has no significant positive correlation with grain nitrogen, but N deficiency can lead to significantly smaller grain size and lower nitrogen content (Dupont & Altenbach, 2003). Grain synthesis can be decomposed into two parts: the rate of synthesis and the duration of the synthesis. For nitrogen grain filling, leaf senescence was found to be negatively correlated to leaves chlorophyll content (Kichey, Hirel, Heumez, Dubois, & Le Gouis, 2007), which can indicate the duration of the nitrogen transfer phase. Positive correlations between the grain yield and the flag leaf photosynthesis have also been reported in the literature (Kichey, Hirel, Heumez, Dubois, & Le Gouis, 2007), (Gaju, et al., 2011) however mostly during stressed growing conditions, such as drought or heatwave. Temperature also impacts the duration and rate of the grain-filling phase, and therefore the grain yield.

Being able to timely identify the activities related to grain nitrogen and protein yield is therefore critical for an accurate estimate and it is important to know the growing stage of the wheat. In this study, we are using the Feekes growth stage systems, which are commonly used, and assigns a numerical code ranging from 1 to 11.4 based on the occurrence of critical growth steps, such as tillering, leaf, and head emergence, and flowering. These stages are summarized in Figure 34 and described in the following section.

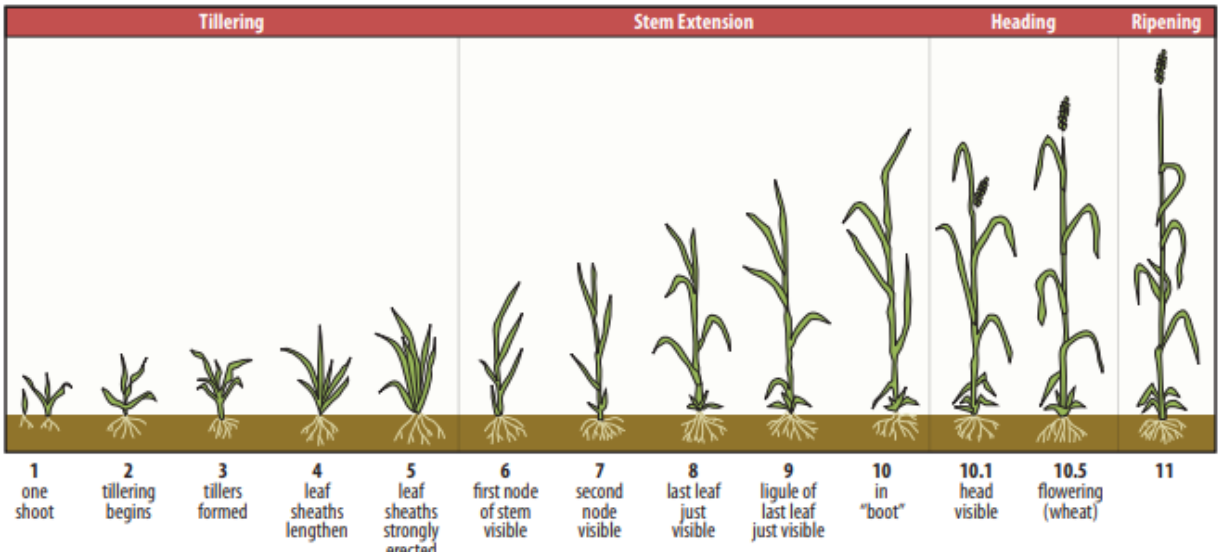

Figure 34 – Feekes' growth stages of wheat (Excerpt from ID-125, A Comprehensive Guide to Wheat Management in Kentucky).

The first five steps can be considered as vegetative stages and required some cool temperature (vernalization): Feekes 1.0 is the emergence of the leaves from the main shoot (the main steps). The tillering begins during Feekes 2.0 (typically in the Fall for the HRW wheat), and the tillers form during Feekes 3.0. Most of the tillers contributing to grain yield are completed at this step. Leaf-sheaths thicken during the fourth step – Feekes 4.0 – (March-April) and strongly erect during Feekes 5.0 (early to mid-April). The first node (visible knot or bump on the tillers) is visible at Feekes 6.0 (mid-April), and the second node becomes visible at Feekes 7.0 (late April to early May). At Feekes 8.0 (Late April to early May), the last leaf (also called flag leaf) begins to emerge from the whorl. Roughly 75% of the effective leaf area contributing to the grain filling phase comes from the flag leaf. This flag leaf fully emerges with a visible ligule during the Feekes 9.0 steps (early May), and the head then starts to develop at Feekes 10.0 or boot stage (mid-May). The head emerges from the head from the leaf sheath of the flag leaf during multiple phases referred

to as Feekes 10.1 to 10.5 (mid to late May). The flowering occurs at Feekes 10.5.1 to 10.5.3 (mid-May to early June). The grain-fill period can take from 13 days to 20 days, depending on the environmental conditions. The ripening and maturation of the kernels happen at Feekes 10.5.4 to 11.4.

Added-value of the Solar-induced chlorophyll fluorescence.

In this study, we are using Solar-induced chlorophyll fluorescence (SIF) observations from satellite imageries along with climatic variables (daily temperature, precipitation) during the grain filling phase of the wheat. Chlorophyll fluorescence is an excellent indicator measuring photosynthesis activity and it can be used as a proxy to measure plant stress due to environmental constraints. As described in the previous section, it is also correlated to the grain filling phase, and therefore an indicator of the nitrogen and protein grain yield.

Figure 35 shows the evolution of SIF activity from 2009 to 2017 over Kansas. It is interesting to note that the year 2011 exhibits a lower than usual average value while the year 2012 and 2013 exhibits an increase to the peak value respectively earlier than usual (2012) and later than usual (2013).

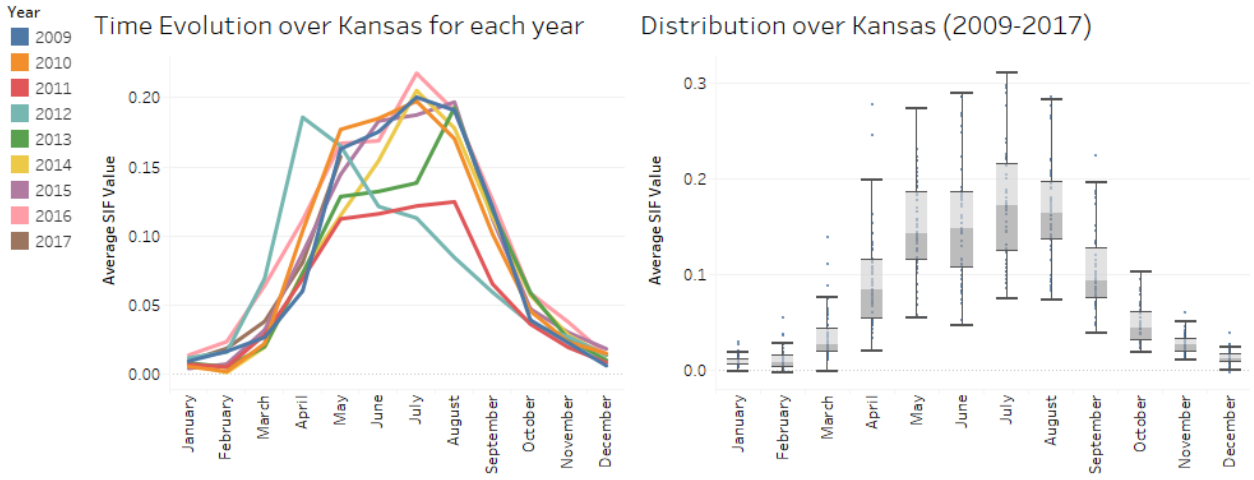

Figure 35 – Temporal evolution and distribution of Sun-Induced Fluorescence values over Kansas. Left: Monthly evolution of SIF from 2009-2017 on Kansas for each year. Right: Average values and confidence intervals of the monthly SIF values over Kansas.

The idea of using vegetation indices in addition to weather indices for protein content estimation is not new. However, most of the research is focusing on using ground-based remote sensing tools, which have the advantages of having an extremely high resolution, full coverage regardless of the cloud cover, and a temporal resolution as high as needed. Canopy reflectance measurement data during the stem extension (Feekes 6.0), booting (Feekes 9.0), and beginning of heading (Feekes 10.1), and several vegetation-derived indices were used to predict winter wheat and barley's protein content at the field scale (Hansen, Jorgensen, & Thomsen, 2002) with mixed results. Using data further into the growing season has produced stronger results, and wheat protein content has been reported to be strongly correlated to NDVI derived rate of ripening (Magney, Eitel, Huggin, & Vierling, 2016), i.e., during the Feekes 11.0 or the month of June for winter wheat in the United States. However, field experiments in Oklahoma using NDVI at several growth

stages of winter wheat (Feekes 8.0 to harvest) concluded that grain N intake, which is a proxy of protein content, could not be reliably predicted using NDVI at any stage of the growth. The Ratio Vegetation Index has also been reported to be highly correlated to the protein content during the grain filling and anthesis stages (Wang, Tian, Yao., Zhu, & Cao, 2014).

These mixed results can be partially explained by the fact that simple vegetation indices alone cannot capture the efficiency at which plants are moving nitrogen to the grain and the losses occurring during this process, an obstacle that the chlorophyll content deduced from the SIF values should better handle.

Move over, simple vegetation indices have been long used as they can be a good indicator of the crop N status, which was assumed to be directly related to the grain protein content (Matre, et al., 2006), (Wang, et al., 2007). Heat or water stresses are also contributing to the rate and efficiency of N uptake to the grain. However, the relationships between protein content and climatic variable are not well understood yet and several regression models using weather parameters such as precipitation, temperature, and sunshine hours are trying to better predict the protein content (Taylor & Gilmour, 1971), (Hansen, Jorgensen, & Thomsen, 2002), (Liu, et al., 2006). Overall, too high temperatures tend to create less numerous (lower yield) but larger and denser kernels (higher protein content) leading to a well-known inverse relationship between yield and protein content. However, these relationships are complex and can vary a lot depending on the plantar used and the climate. Field experiments in Sweden showed that temperature during the growing season (September to August) was the major parameter, explaining 34% of the variance in

protein concentration in spring wheat, and 71% for winter wheat (Johansson & Svensson, 1998) and that precipitation during the last months positively impacted the protein content. On the other hand, high precipitation one to two months before the maturity of the plants and temperature two to three weeks before the maturity has been reported to negatively impact protein content in Southwestern Nebraska (Smika & Greb, 1973). Another field-based study in China showed that the protein content is highly correlated to the range of daily temperature (Tmin – Tmax) during the growing seasons for high-protein cultivars and to the average temperature and total sunshine hours for medium-protein cultivar (Pan, Zhu, Cao, Dai, & Jiang, 2006). Durum wheat, which is higher protein content, has been reported to have protein content positively correlated with air temperature during the growing stage and negatively correlated with precipitation in 4 sites in Italy, where it is one major class of wheat cultivated (Dalla Marta, Grifoni, Mancini, Zipoli, & Orlandini, 2011).

For these reasons, we are using precipitation, maximum daily temperature, and Sun-Induced Chlorophyll Fluorescence as variables to predict the grain's protein content. As we aim to build a predictive tool, we excluded using yield, which is only accessible after the harvest or can be predicted but would add a layer of uncertainty to our model.

Data

We developed a code to extract data from the "Hard Red Winter Wheat Regional Quality Survey" annual reports from the Plain Grain Inc. which performed laboratory analysis of the production on the United States in collaboration with the USDA/ARS and Hard Winter Wheat Quality Lab. The data are reported at the grainshed scale, i.e., several

counties gathered. Figure 36 shows the data available from 2000 to 2017 from the Plain Grain Inc. and the grainshed division used.

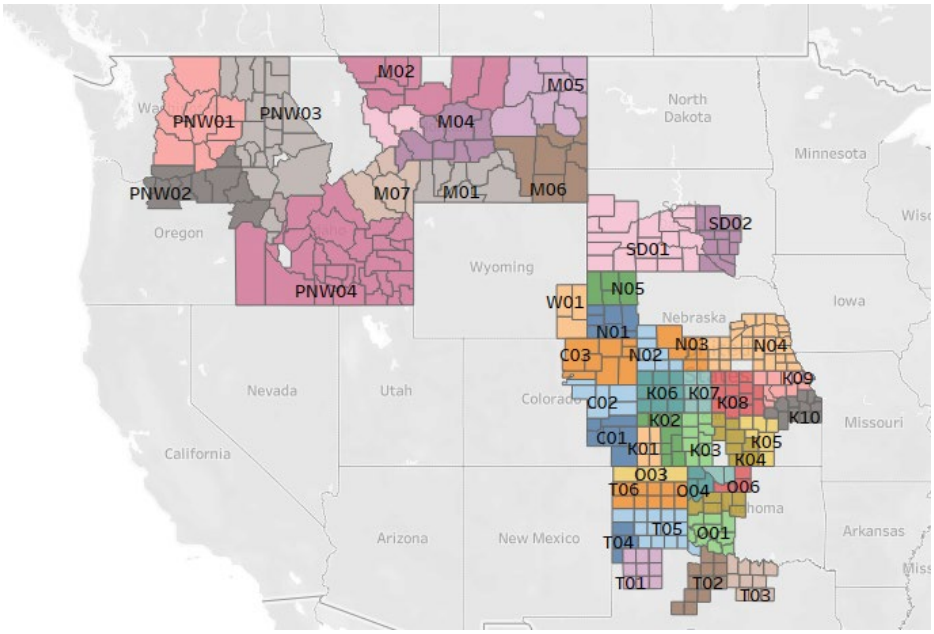

Figure 36 – Data availability by agricultural districts on Hard Red Winter Wheat.

Unfortunately, these grainsheds are different from USDA agricultural districts. Moreover, the nomenclature changed in 2018 in Kansas, and grainshed have been rearranged. We have nevertheless one unique dataset of valuable quantitative data on the quality of HRW Wheat from 2009 to 2017 and from 2018 to 2020.

We used only the data until 2017 for the training/validation of our model. We matched grainsheds with USDA NASS's counties to add production, area, and calculated yields at grainshed scale and incorporated them on the Google Earth Engine platform. In addition to these tabular data, we used remotely sensed raster data, including precipitation and temperature from PRISM, and Solar-induced chlorophyll fluorescence from Li et al. (Li &

Xiao, A global, 0.05-degree product of solar-induced chlorophyll fluorescence, 2019). The SIF dataset is at relatively high resolution (0.05 degree), 8 days temporal resolution from 2000 to now. Climate and phonologic data were extracted at county scale using Google Earth Engine.

Method

Several models were built using different combinations of variables. We calculate average maximum temperature, average precipitation, and average SIF for several months ahead of the harvest, and then derived the monthly anomalies (departure from average at grainshed scale). For the anomaly calculations, we used a base period from 2000 to 2020 (all the data available) and we used the same base period for temperature and precipitation for consistency. Because the range of protein content is more important than the actual value, we decided to use a support vector machine's-based regression model, which fits a model and keeps the error to a certain margin by reducing the coefficient vector rather than the squared error. We used the scikit-learn library in Python to fit the regression models using a support vector machine (SVM) regression model, with a radial basis function (RBF) and using 70% of the dataset for training and 30% for testing. We computed the determination coefficient – R^2 – between the observed and predicted protein content for both training and testing and performed a repeated random sub-sampling validation, as a robustness check, using the same 70/30 ratio for training and testing. Comparing the determination coefficient between the training and the testing informs of potential overfitting issues with the model. More information on the model can be found in Appendix 3.2. Confidence intervals of the determination coefficient were calculated using

100 repeated tests. Validations of the results were also visualized using the class of wheat protein content, using three categories – i.e., low, medium, high – respectively defined by the 25th, 50th, and 75th percentiles since it is one of the criteria to grade wheat. Table 5 summarizes the observed percentiles in protein content for wheat. We used PGI's protein content at grainshed division, and weighted-average the protein content using the total production obtained from USDA NASS such as:

$$PC[g/100g] = \frac{\sum_{PGI} PC_{PGI} \cdot P_{PGI}}{P_{PGI}}$$

Where PC is the protein content in [g/100g] and P is the total HRW wheat production in tonnes.

Table 5 – 25th, 50th, and 75th percentiles protein content of wheat.

Class	**Wheat Protein content [g/100g]**
Low	9.50 – 11.60 (10.55)
Middle	11.60 – 13.40 (12.50)
High	13.40 – 16.90 (15.15)

Results

We selected the protein content model with the best performance and the results are displayed in Figure 37. Models with good performance are in a darker color. The best results are obtained using the Sun-Induced Fluorescence and temperatures anomalies in May (one to two months before the beginning of the harvest). The disparity between the training and testing correlation coefficients in March, June, and July suggests that the model

is overfitting the data. June and July exhibit very poor performances, which was expected as the harvest is typically happening from Mid-June to the end of July.

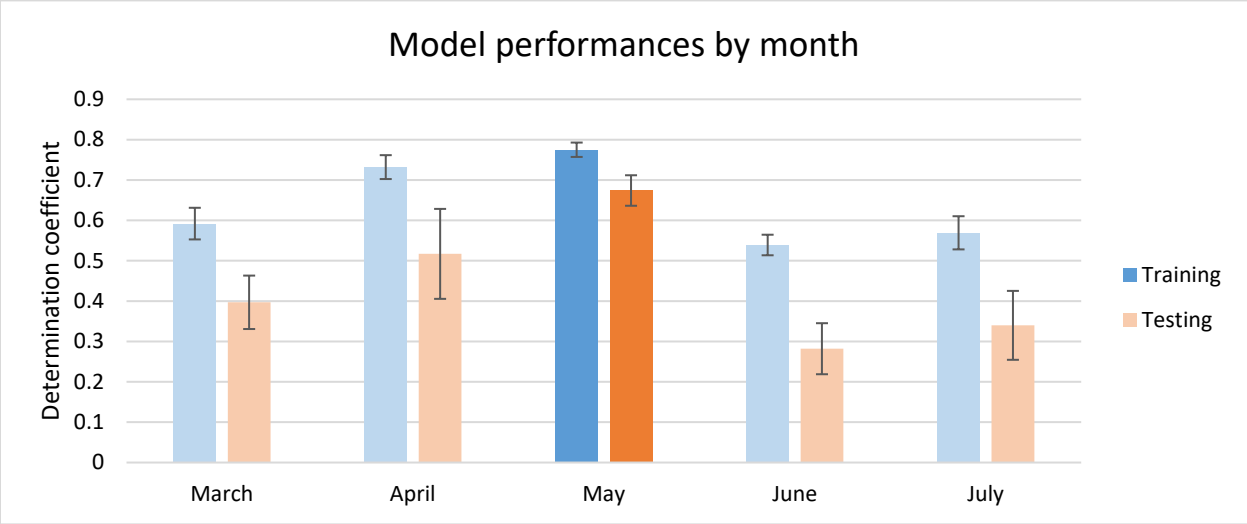

Figure 37 – Model performances by month ahead of the harvest. Darker colors indicate a model with a good performance.

Figure 38 summarizes the prediction performance of the best model, looking at the class of protein content rather than actual values. Overall, the model accurately classifies the ranges of protein content, with better results for medium values.

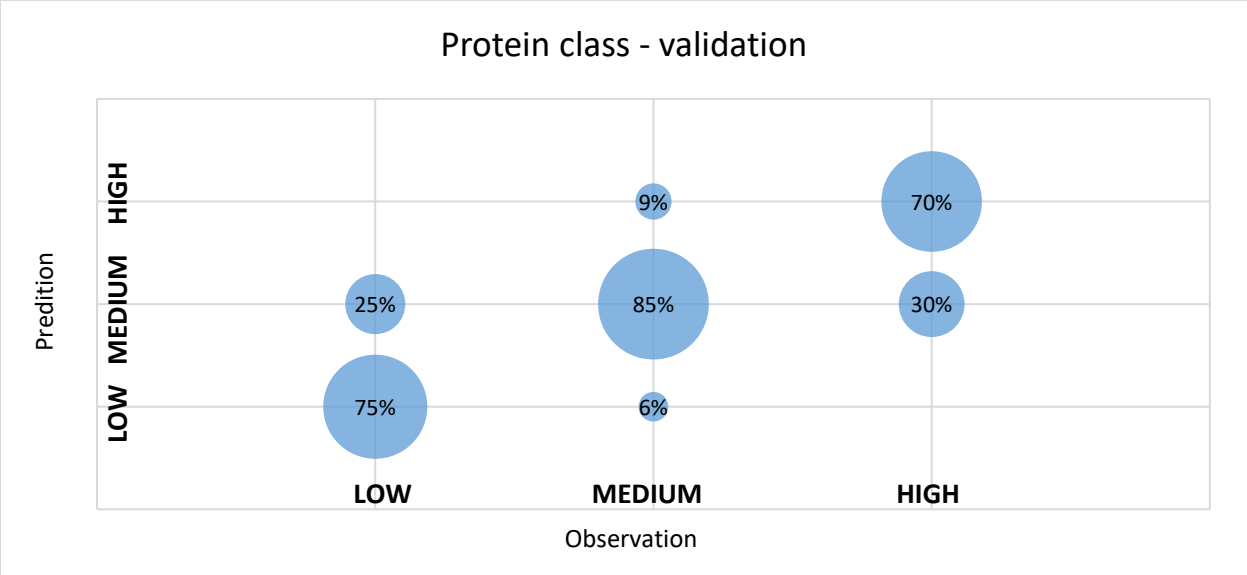

Figure 38 – Validations of the wheat classification based on the protein content's value.

Discussion

The best performances are obtained by the model using the May data, which corresponds to the growing stages of Feekes 8.0 – the emergence of the flag leaf – to Feekes 10.5 – heading and booting. This result was expected as the flag leaf account for roughly 75% of the effective lead area contributing to the grain filling phase and the nitrogen transport from the vegetative organs to the grain happens during this phase. Using June and July data is leading to poor performance of the model and significant difference between the training and testing determination coefficients, suggesting no relationship. Using March or April data also lead to poor performances of the model suggesting no correlation between the March and April months to the grain protein content. Although also important, these two months correspond to the jointing of the wheat and start of heading, so the pre-anthesis nitrogen uptake by the plant and not to the transport of N from the stem and leaf to the grain. Our findings are strengthened by other studies using remote sensing data acquired one month before the harvest to predict the grain protein

content (Ozturk & Aydin, 2004), (Liu, et al., 2006), (Li-Hong, Wei-Xing, & Lin-Zhang, 2007), (Wang, Tian, Yao., Zhu, & Cao, 2014). However, most of the other studies are focusing on measuring the nitrogen status of the wheat using one or several vegetation indices, such as the NDVI – Normalized difference vegetation index, the SIPI – Structure Intensive Pigment Index, the PSRI – Plant Senescence Reflectance Index, and the EVI – Enhanced Vegetation Index. These indices, although able to monitor the Leaf Nitrogen Content (LNC), are reported not to capture entirely the heat and nitrogen stresses, which are limiting factors for protein synthesis (Jia, Rossini, Colombo, & Celesti, 2021). Sun-Induced Fluorescence, on the other hand, is related to the plant physiology rate and can better capture these stressors (Burling, Hunsche, & Noga, 2011). In another study (Song, et al., 2018), Sun-Induced Fluorescence values were reported to reflect almost instantaneously the effects of the heat stress while other vegetations indices, such as NDVI and EVI, lagged nearly one month.

In this study, we used eight years of data, which include a major drought, to train and test of model using solely remotely sensed data. While most research has looked at predicting grains' protein content using vegetation indices, most of the studies are using one year of data, and there is no guarantee that the model developed would work another year with different levels of water and heat stresses which impact dramatically the physiological processes of the plant.

By using averaging weekly data to monthly data, we minimize the error of selecting a time corresponding to the growing stage of interests. However, growth stages can vary from year to year, sometimes significantly and some long-term trends in the beginning and

end of some critical crop operations have been detected. For instance, the planting season of HRW wheat is delayed by one week every decade in most states since 1981. Moreover, for instance, 2012's heading stage happened two weeks earlier than the average and the harvest 3 weeks earlier than the average in Kansas. However, we are losing the 8-days resolution of the SIF dataset. There is a great potential of using SIF or vegetation indices to derive the actual growing stages of the wheat (Magney, Eitel, Huggin, & Vierling, 2016) and use the weekly SIF data to their full potential. Appendix 3.1. provides a more comprehensive analysis of these long-term trends and short-term variations.

Conclusion

We presented a regression model based on a supervised classification to forecast the hard red winter wheat's protein content at the grainshed division scale using solely remotely sensed data. We extracted HRW wheat's protein content from harvest reports conducted from 2009 to 2017 in 43 US subdivisions accounting for more than 95% of the total HRW wheat production in the United States. We extracted Sun-Induced Chlorophyll Fluorescence data from GOSIF and daily maximum temperature from PRISM to train and validate our model. Our dataset is covering twelve US states, including several states and years under severe drought and heat stress, and is representative of several growing conditions. Our results show that Sun-Induced Chlorophyll Fluorescence can well capture the physiological processes behind nitrogen transport to the grain and protein synthesis. Our model can be used to predict grain protein content in HRW wheat with one to two months lead time and it can correctly classify grain protein content between low, medium, and high values.

In our study, we average our SIF data to monthly data. There are strong opportunities to use SIF combined with other vegetation indices to derive the actual growing stages of the wheat instead of using a crop calendar and to use the 8-days SIF data for more precise monitoring of the predicting grain protein content.

CHAPTER 4

MAPPING RICE PADDIES AREAS USING REMOTE SENSING SATELLITE DATA AND SUPPORT VECTOR MACHINE ALGORITHM.

Introduction

Rice is a critical crop, representing around 11% of the global crop area and 20% of the dietary energy supply for more than half of the population. With increasing food demand and population growth, rice is expected to stay a major stable crop and a major source of protein for most of the planet. It is mostly locally grown in Asia, imported in the Middle East and Oceania, and equally locally produced and imported in West Africa. Figure 39 summarizes the global importance of rice to the food plate.

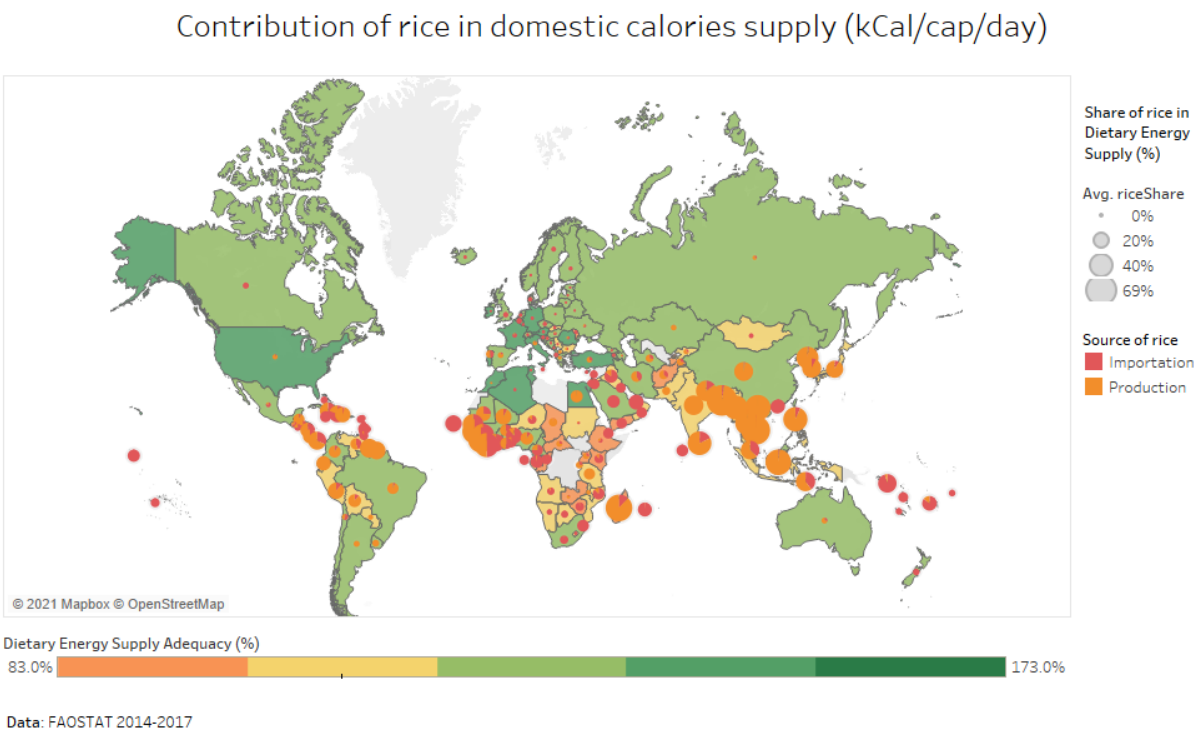

Figure 39 – Global analysis of the importance of rice to the food plate. The background color indicates the percentage adequacy in dietary energy supply. The pie's size represents the share of rice to the country's calory supply, and the colors indicate the origin of the rice (locally produced or imported).

Mapping paddy rice is an important step for rice production management, for better forecasting rice production, improving the information on rice-growing systems, and crops conditions. Thanks to higher yield obtained from fertilizers, higher water resources availability, and better pest control, rice production was able to increase within the past decades at roughly constant planted area. Figure 40 reveals the temporal evolutions of yield and planted rice area at the country scale from 1961 to 2019. While most of the major producers increased their yield much faster than their planted area, India is a noticeable exception and has an increased its yield much less than other countries and has strong variation in planted area. There are many causes of changing of planted area: better management techniques resulting in higher yield and lower land requirements, conversions from rice paddies to other crops, which can be higher values, or less water intense, local policies, frequent droughts can lead to long-term increasing or decreasing trends in planted area, but also having a third growing season of rice. Local events, such as cold fronts delaying or canceling the planting season, extreme droughts, extreme monsoon can all result in temporary decreased planted or harvested area.

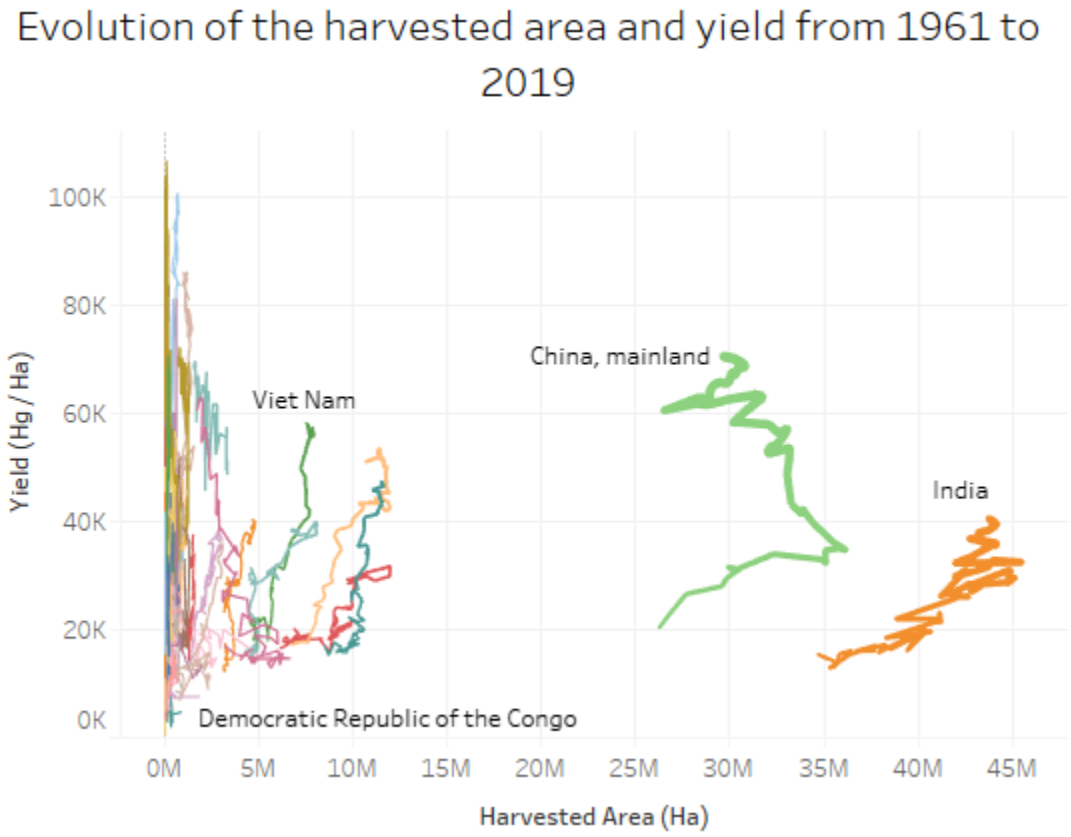

Figure 40 – Evolution of the harvested area and rice yield. The line is following the time series from 1961 to 2019, the thickness indicates the total production of the country, and the color indicates the country. Data come from FAOSTAT.

Current statistics are reported by farmers after the harvest or by governments after in-situ sampling. Landcover land-use datasets are typically updated every decade, national statistics are updated long after the harvesting time and compiled on FAO. There is therefore a strong need to be able to map rice paddies from planting to harvesting to be able to capture these changes. Moreover, most of the statistics are provided by year and not growing seasons. While North America and Europe typically have one growing season, most of the Asian countries have two or three. Bangladesh for instance has three growing

seasons, where the main harvest happens in November for *Aman* rice (rainy season), the secondary harvests happen in April-May for the *Boro* rice (Dry season) and in July-August for the *Aus* rice. China, India, (two main producers) typically have 2 growing seasons and some parts can have up to three as well. In general, the last growing season can be optional if the two previous ones have been delayed and there is no time to seed, grow, and harvest the rice before the beginning of the next growing season, which is the main one. In summary, there is a strong need to have harvested data by growing season instead of the total harvested area by year. Moreover, in South Asia, wheat is commonly planted after rice or double-planted with rice, therefore having information on rice area can also serve to better estimate wheat production. In Pakistan for instance half of the wheat is planted after rice or cotton (Sheema, Ali, & Akhtar, 2002). In other parts, wheat can be double planted with rice, often time after the optimal planting period of wheat as it is only the secondary product (Hobbs & Mehla, 2003). As having three seasons in a year requires planning, favorable meteorological conditions, and timely seeding and harvesting of the first two seasons, it is important to be able to monitor the start and end of the growing season as well.

A few rice crops calendar already exist, for instance from the FAO (Food and Agriculture Organization, 2021), from Sacks et al. (Sacks, Deryng, Foley, & Ramankutty, 2010), the MIRCA2000 (Portmann, Siebert, & Döll, 2010), and RiceAtlas (Laborte, et al., 2017) to name a few. While the FAO dataset has the longest time record on a global scale, it has limited and outdated information. The dataset from Sacks et al. provides planting and harvesting dates for 19 major crops at the national or subnational level by compiling different sources of data, mainly from USDA, and AGRIMET. By combining different

datasets, it can increase the level of detail but still inherits poor coverage in developing countries and does not capture change in times year after year. MIRCA2000 is similar as they are using the same primary sources; it is proving a gridded calendar for rainfed and irrigated rice separately but fails to fully capture rice lands that are cultivated more than once a year. Finally, RiceAtlas is focusing on the rice calendar and provides information at a higher spatial resolution (finer subregions, especially in Asia) and capture more rice lands with more than one growing season but still inherits of the limitations of the source of data, which are statistics and not direct observations.

Being able to detect the number of the growing season and to provide a start and end of the growing season at a high resolution is therefore important to better analyze rice lands, change in growing seasons, and change in growth patterns.

Many models, physical-processed based or data-driven, are indeed focusing on predicting rice yield to then infer the production. Satellite imageries have been widely used for such purposes (Casanova, Goudriaan, Catala Forner, & Withagen, 2002), (Huang, Wang, Li, Tian, & Pan, 2013) and models typically exhibit good performances. While changes in yield are significantly impacting the change in production, changes in areas are also explaining a significant part of the annual variations. Figure 41 shows the percentage change of harvested area and yield (using detrended data) for major rice producers.

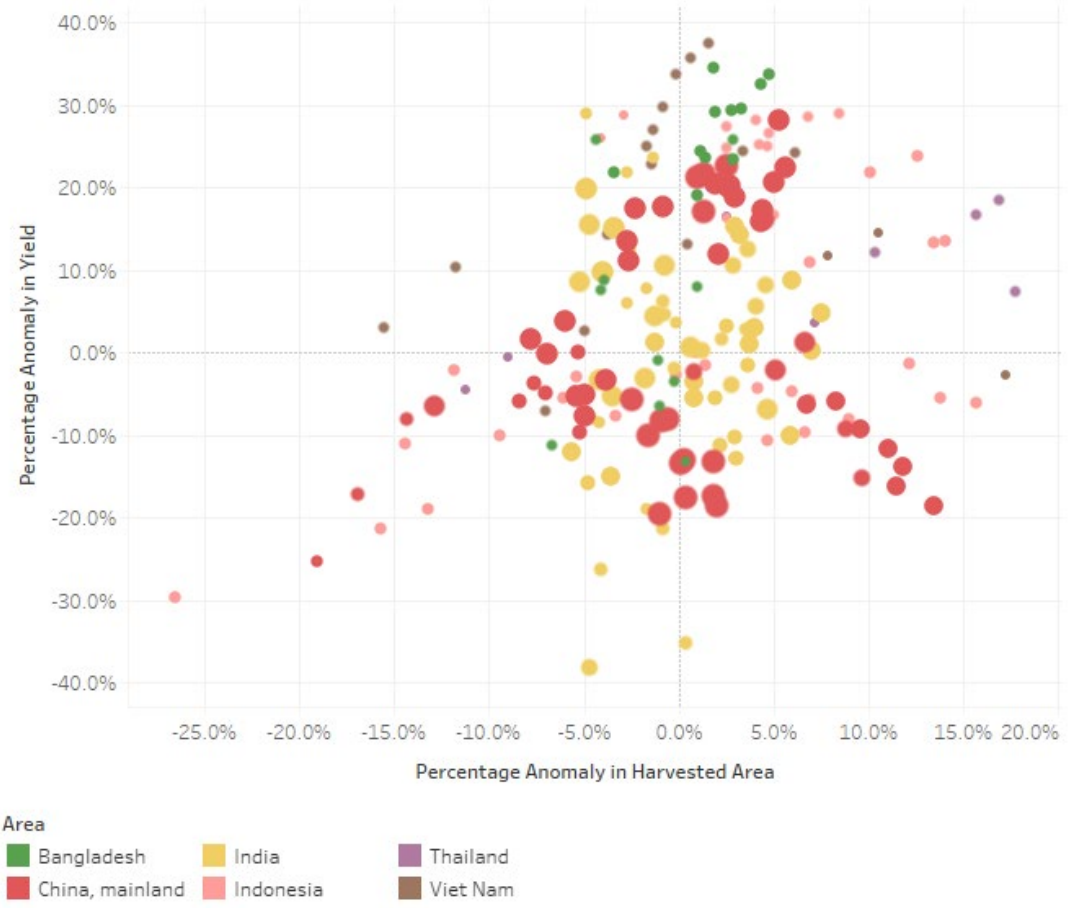

Figure 41 – Change in yield and harvested area for the main rice producers. The size of the circle is proportional to the total production and the color indicates the country. Data come from FAOSTAT.

The total change in production was then decoupled using the formula below and one example for Bangladesh is represented in Figure 43:

$$\Delta P = P_n - P_{n-1}$$

$$\Delta P = (A_n \cdot Y_n) - (A_{n-1} \cdot Y_{n-1})$$

$$\Delta P = (A_n - A_{n-1}) \cdot Y_n + A_{n-1} \cdot (Y_n - Y_{n-1})$$

$$\Delta P = \Delta A \cdot Y_n + A_{n-1} \cdot \Delta Y$$

Figure 42 shows that changes in harvested areas can contribute to a loss (or increase) of rice production. In some cases, the gains in yields are annihilated by the loss of harvested areas (e.g. the years 1983, 2016, 2018).

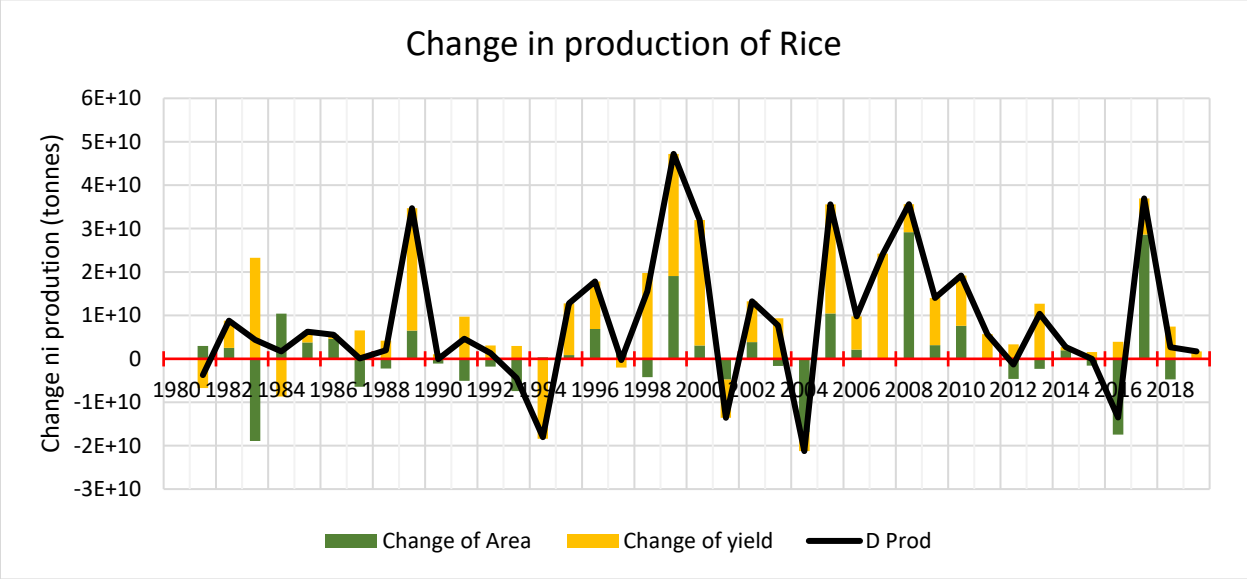

Figure 42 – Change in rice production in Bangladesh and contributions of the change in yield and change in areas. Data come from FAOSTAT.

In addition to rice production, mapping rice paddies can have a considerable advantage to better estimate greenhouse gases emission. Growing rice in flooding conditions is causing the paddy soil to become temporary anoxic, i.e., depleted of dissolved oxygen. In such conditions, bacteria are producing methane (Yao & Conrad, 1999). As 90% of the rice land is temporarily flooded, such methane emissions are not negligible and rice paddies have already been estimated to be responsible for 12% of the global emissions of methane, and 3% of the global anthropogenic warming (Spahni, et al., 2011), (Yan, Akiyama, Yagi, & Akmoto, 2009).

Being able to accurately map the surface area used for rice paddies is therefore crucial for being able to predict rice production and to better monitor CH4 emissions.

Knowledge of the growth stages of rice is essential to correctly differentiate rice paddies from other vegetation or temporary flooded area and to assess the uncertainties of this study. The growing cycle can be summarized in three main phases: the vegetative phase, the reproductive phase, and the ripening phase. The last two phases last for about 30 days and the total growing cycle can take from three to six months depending on the cultivar used, and the local growing conditions (Yoshida, 1981), (Torbick, Chowdhury, Salas, & Qi, 2017). Furthermore, the crop calendar mainly depends on the water distribution system, meaning that it can change from crop to crop within the same neighborhood (Le Toan, et al., 1997), (Alauddin & Sharma, 2013). In the countries investigated in this study, rice is usually grown in two to three seasons. We will refer to winter rice as “Aman”, summer rice as “Boro”, and autumn rice as “Aus”. The Boro rice is typically planted from December to February and harvested from April to May and is immediately followed by the Aman rice harvested from November to December or by the Aus rice harvested from July to August. Figure 43 summarizes the growing steps for transplanted rice, which is the most common and elaborative method of crop establishment for rice in Asia (IRRI, 2021).

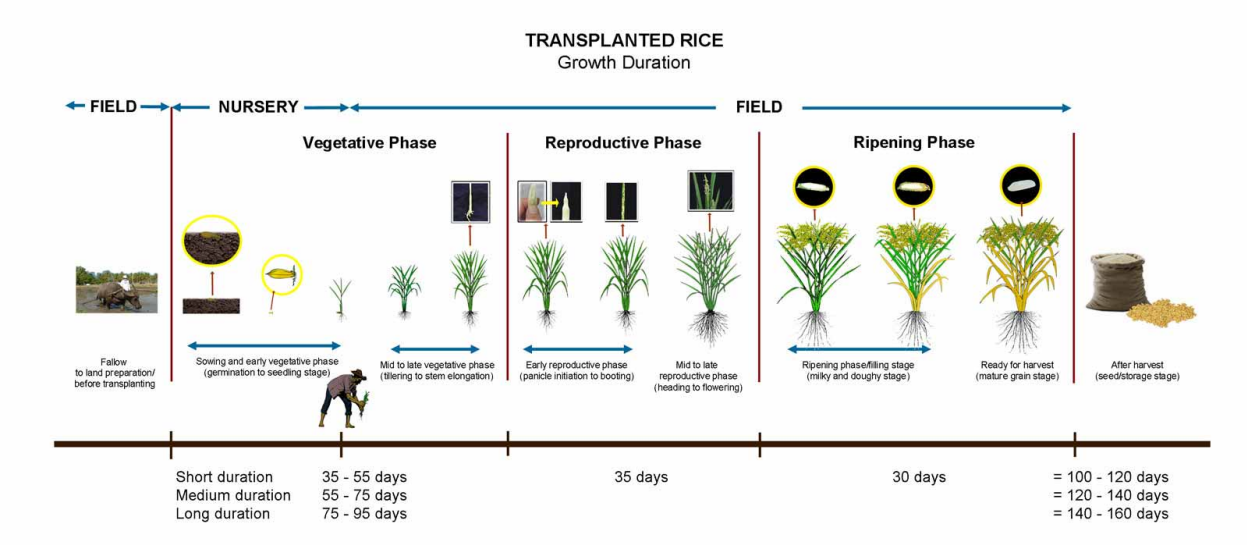

Figure 43 – Growing phase of transplanted rice - Source: International Rice Research Institute (http://www.knowledgebank.irri.org/step-by-step-production/pre-planting/crop-calendar)

Crop patterns are also critical to better calibrate models and interpret uncertainties in the classification. Figure 44 summarizes the distribution of crop patterns obtained from (Nasim, et al., 2017). Roughly half of the total cropped area consists in one or two growing seasons of rice (27% consists in Boro rice only, 6% in Aman rice only, and 13% in both). The third growing season of Aus rice is rarer (3% for Boro-Aus-Aman). In this study, we will focus solely on the Boro and Aman patterns.

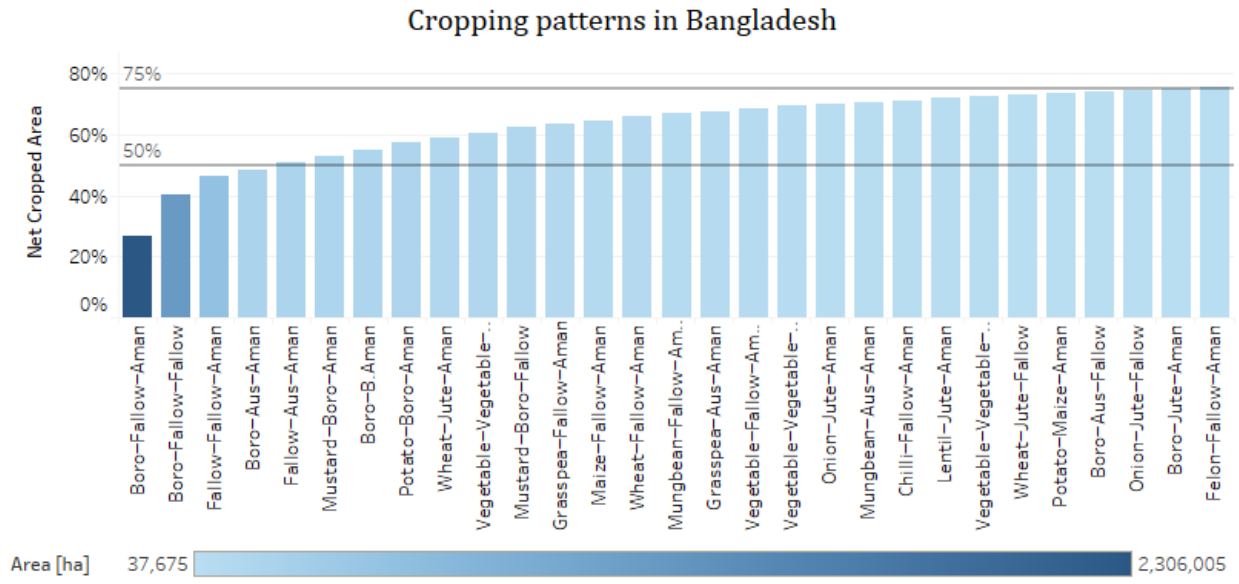

Figure 44 – Main crop patterns in Bangladesh in 2015. Share of each crop pattern to the total cropped area. Data come from (Nasim, et al., 2017). Boro, Aus, and Aman refer to the three possible growing seasons of rice.

Mapping rice paddies got easier with the advancement of higher resolution satellite imageries and the development of new machine learning tools. Optical sensors from Landsat and NOAA Advanced Very High-Resolution Radiometer (AVHRR) have been used to classify rice lands using the fine-resolution Landsat Thematic Mapper (TM) or the time evolution of the moderate-resolution AVHRR's derived Normalized Difference Vegetation Index (NDVI) (Fang, 1988), (Okamoto & Fukuhara, 1996), (Van Niel, McVicar, Fang, & Liang, 2003). The Moderate Resolution Imaging Spectroradiometer (MODIS) has been widely used thanks to its visible and near-infrared bands from which we can calculate several vegetation indices such as the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Land Surface Water Index (LSWI) (Xiao, et al., 2005), (Peng, Huete, Huang, Wang, & Sun, 2011). The LSWI is generally used to detect changes in

surface moisture during the flooding and transplanting phases. However, most rice paddies are in rainy areas, with large cloud covers, making optical sensors from satellites not suitable for such analysis and for monitoring monthly rice crops progress. On the other side, microwave sensors however can penetrate clouds and the Copernicus Sentinel-1 operated from the European Space Agency provides co- and cross-polarized backscatter from its Synthetic Aperture Radar (SAR). It enables the detection of flooded areas at a high spatial and temporal resolution. Several studies have been using SAR sensor data from Landsat, MODIS, and ALOS PALSAR for rice mapping (Park, et al., 2018), (Xie, Xhang, Wu, Wang, & Zhang, 2015), typically using the co-polarization ratio (HH/HV). Cross-polarized backscatter VH is sensitive to rice crops as it captures well the temporary flooding during the transplantation phase (see Appendix 4.1. for more details).

Support Vector Machine (SVM) and Random Forest (RF) are popular classification algorithms developed in the early 90s (Boser, Guyon, & Vapnik, 1992). They have long been used for land cover classification (Brown, Gunn, & Lewis, 1999), (Fukuda & Hirosawa, 2001), (Mathur & Foody, 2004) and more recently for rice mapping (Turner & Congalton, 1998), (Tan, Koay, Lim, Ewe, & Chuah, 2007), (Zhang, Wang, Wu, Qi, & Salas, 2009) and exhibit strong classification skills, even with a relatively small training sample. While these two methods can produce results with similar performance, there are fundamental differences between the two approaches. RFs are inherently multiclass classifiers using multiple decision trees and merging the results to build a more stable and accurate classification. SVMs on the other hand, are designed to separate a dataset into two classes by maximizing the margins between the classes' closest points, or in other words, to maximize the width of a hyperplane separating the two classes. Furthermore, RFs tend to

create segmented decision boundaries (i.e. parallel to the x and y-axis in case of a 2D classification) while SVMs can produce much smoother results depending on the kernel used. Previous studies have reported similar accuracy levels using either of the methods. Several studies in relatively small areas and unique years, using randomly distributed training points have shown greater skills for RF (Cai, Lin, & Zhang, 2019) or SVM (Mansaray, Wang, Huang, Yang, & Kanu, 2020). These studies are however typically looking at a small area, with randomly distributed training points all over the area of interest and using training/testing in one or several years.

Validation of the results remains a challenge, mainly because there are no "ground truth" observation datasets of rice lands at high resolution, and incorporating one, two, or three growing seasons. Most of the studies are typically focused on measuring the accuracy of the classification algorithm using their confusion matrix to assess their consistency in the classification. Results are then roughly validated using country average statistics, which encompasses the total rice area for the one, two, or three growing seasons, and using higher resolution county scale in the United States, which has only one growing season. More detailed validation data can also be obtained from harvest reports for each country when available. In this chapter, we created a database of harvest areas using all harvest reports available from official sources. We focused on harvest reports from Bangladesh (Bangladesh Bureau of Statistics, 2021), which provides detailed harvest reports at district and province level for all rice growing seasons, and from 2015 to 2018, which provide two years overlap with the sentinel-1 data used. Other sources of validation include evaluating the model performances and consistency.

Data

In this chapter, we used the Sentinel 1A Interferometric Wide (IW) Ground Range Detected High Resolution (GRDH) product and retrieved the cross-polarized VH backscatter and the VV polarization at 10 meters spatial resolutions and 16 days revisit time. Data were obtained from the Google Earth Engine platform, at the native 10-meters resolution, and log-scaled. Model's results were extracted at the district level using the administrative divisions obtained from the Global Administrative Unit Layer (GAUL) produced by the Food and Agriculture Organization (FAO) in 2014. In this dataset, the province of Mymensingh is incorporated into the province of Dhaka, although they are distinct provinces since 2015.

We also developed a code to extract data from pdf images from harvest reported generated by the Bangladesh Bureau of Statistics (Bangladesh Bureau of Statistics, 2021). We extracted harvested areas, yield, and production in Bangladesh at the district level (using all 64 districts of the 8 provinces) and for the three growing seasons of rice (Aman, Boro, and Aus) as shown on Figure 45.

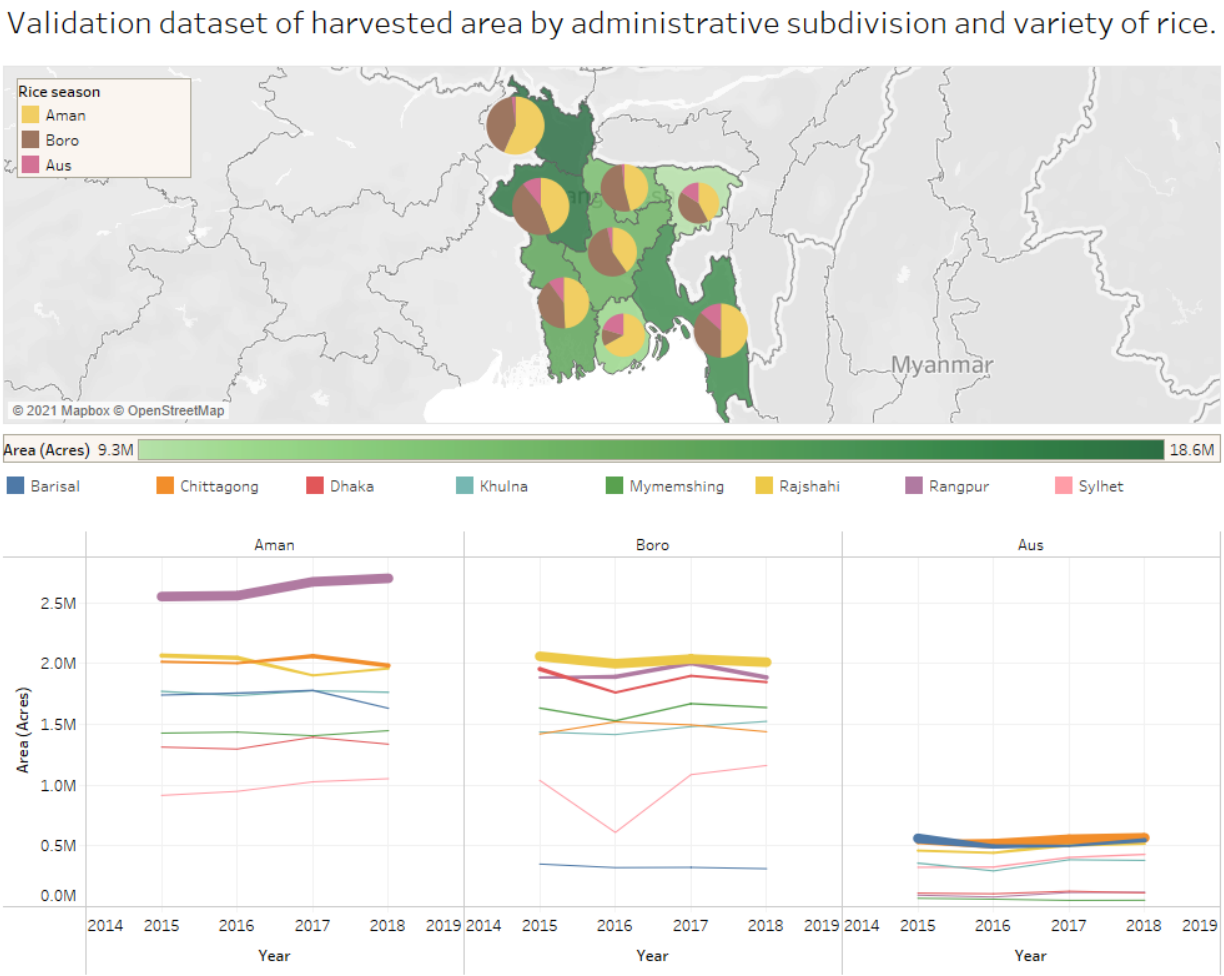

Figure 45 – Ground observations of harvested areas by province and season. Top chart: States visualized by the total harvested area during the three growing seasons. The pie charts indicate the share of each season to the total harvested area. Bottom graph: Harvested areas for each state and growing season. The thickness of the line indicates the percentage of the province to the country's total harvested area.

Figure 45 displays the observation data gathered from the Bangladesh Bureau of Statistics and used for the optimization of the parameters of the SVMs' kernel and validation. The bottom charts reveal the total harvested area in acres during the three main rice-growing seasons (Aman, Boro, and Aus) for all provinces of Bangladesh and from 2015/16 to

2018/19. Financial Year starts on July 1st in Bangladesh, so the area harvested in 2015/16 correspond to the Aman rice planted around 2015 and harvested at the end of 2015, to the Boro rice planted in December 2015 and harvested in Mach to May 2016, and we assumed to the Aus rice planted in March 2015 (previous financial year) and harvested in July 2015.

Method

We set up the SVM classifier on Google Earth Engine, using a Radial Basis Function (RBF), which is widely used, and the Random Forest classifier. We restricted the SAR dataset to the period of interest depending on the growing season of the rice (e.g., July to October for the "Aman" rice in Bangladesh, which represent roughly 40% of the total rice production), manually label several types of landcover, and train the classifier using these labels to classify other areas as shown on Figure 46. We then calculate the total planted/harvested area. We trained the classifiers using several sets of training data, manually entered, for more robust classification and for calculating the confidence intervals of the area.

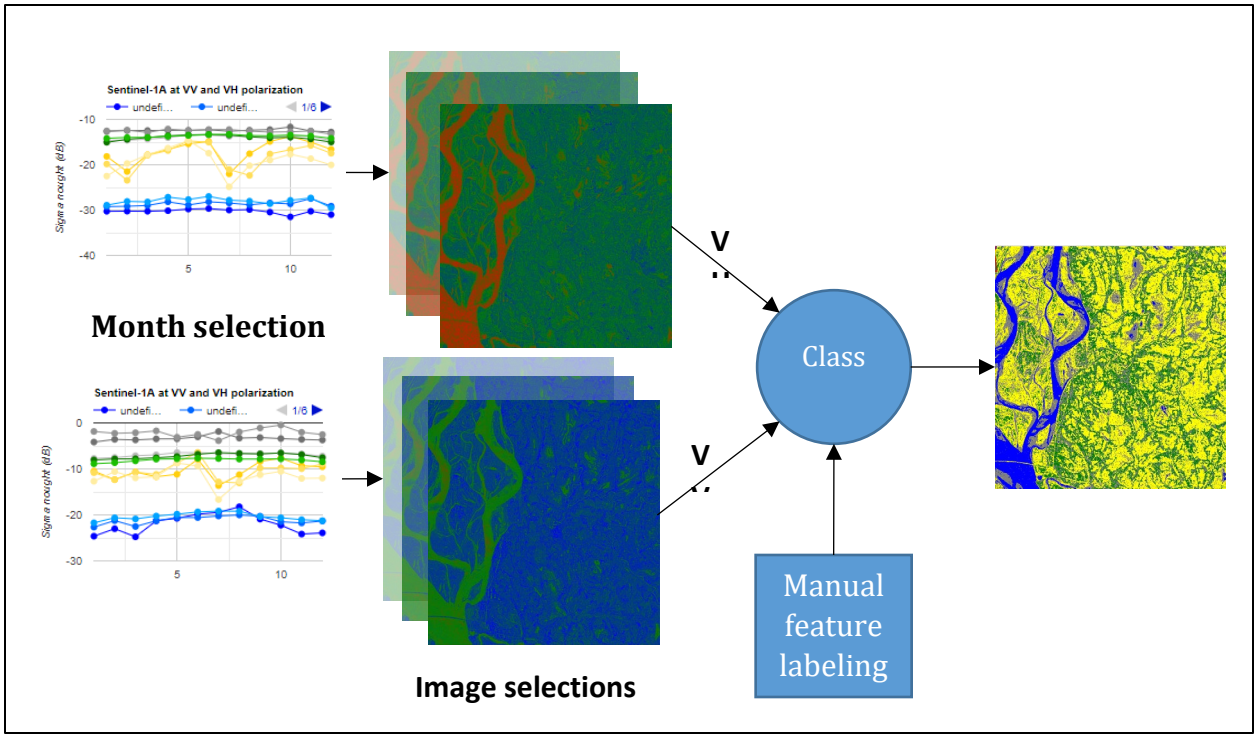

Figure 46 – Workflow of the rice paddies classification. Workflow developed to use Sentinel-1 VV and VH polarizations and a Support Vector Machine Classifier to map several varieties of rice paddies at the provincial level.

Figure 47 displays the SAR signals between several types of land use. Rice areas are temporarily flooded, which is detected by the SAR dataset, while other areas remain relatively constant. Additional datasets were used as a comparison, such as the MODIS EVI and NDVI.

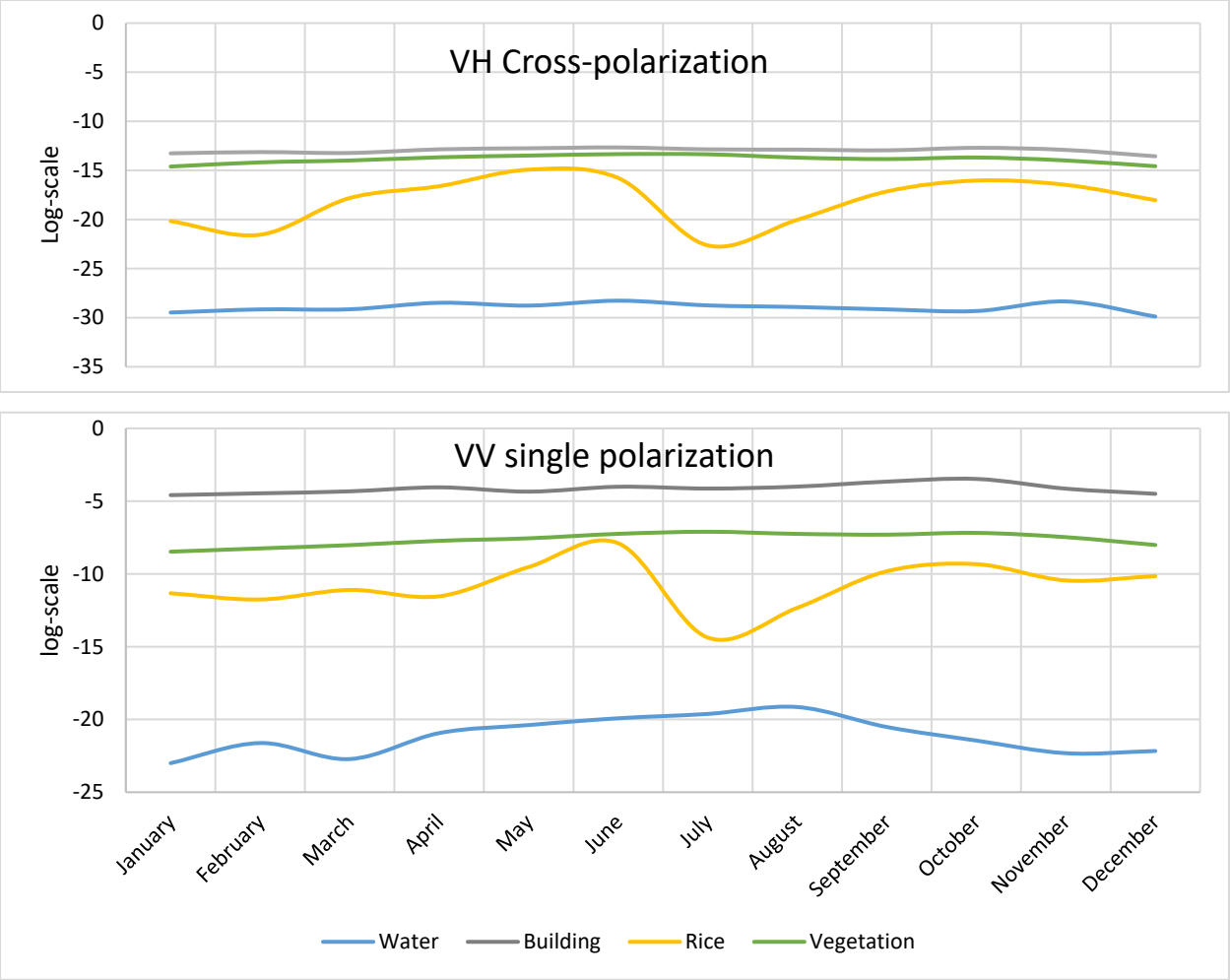

Figure 47 – Monthly evolution of the VH and VV polarizations (Log-scale) on several landcover (water body, building, rice paddies, and other vegetation).

The optimization of the model parameters was done by visually inspecting several randomly selected areas, and by calculating the variance of the total rice areas using different training datasets. The cost and gamma coefficients of the SVM and the number of trees of the RF were optimized using a grid-search approach (Joachims, 2002). For each model, we used several randomly selected training polygons and a series of parameter values. We looked at the variance in total area for each class and each season to choose the optimal parameters. Figure 48 for instance reveals the impact of the cost parameter on the

classification, a higher cost parameter, meaning a larger window used in the kernel function, provides a smoother classification, resulting in a temporary flooded area to be classified as a water body.

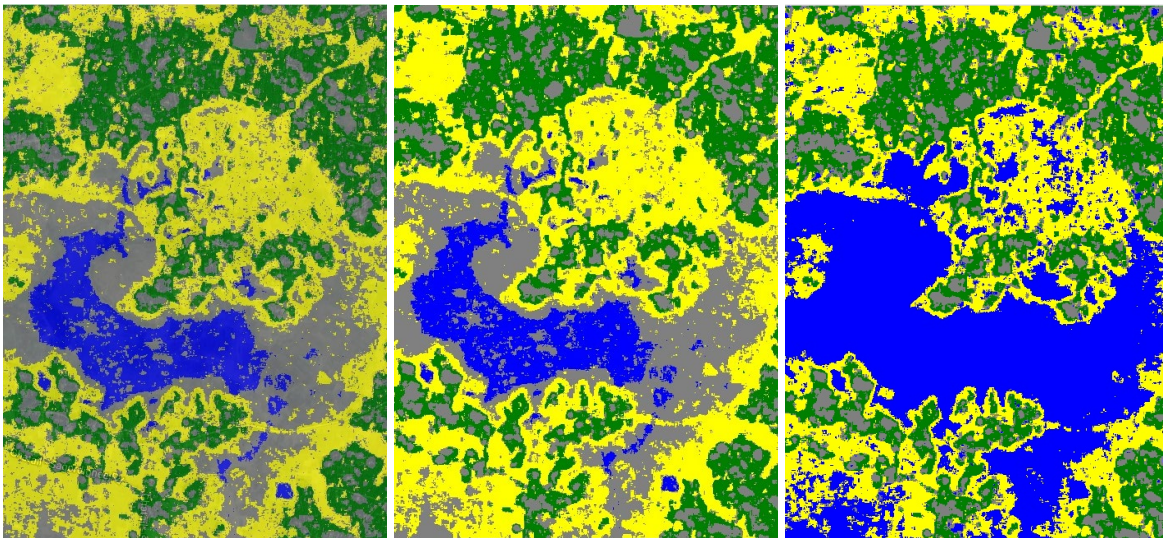

Figure 48 – Classification for several cost parameters. C=10, 1, 0.1 respectively on the left, middle, and right plots.

We selected the parameters with the most consistent prediction skills using building area, which we assumed must be relatively constant over time for each region. Figure 73 summarizes the results using several cost parameters in the SVM algorithm.

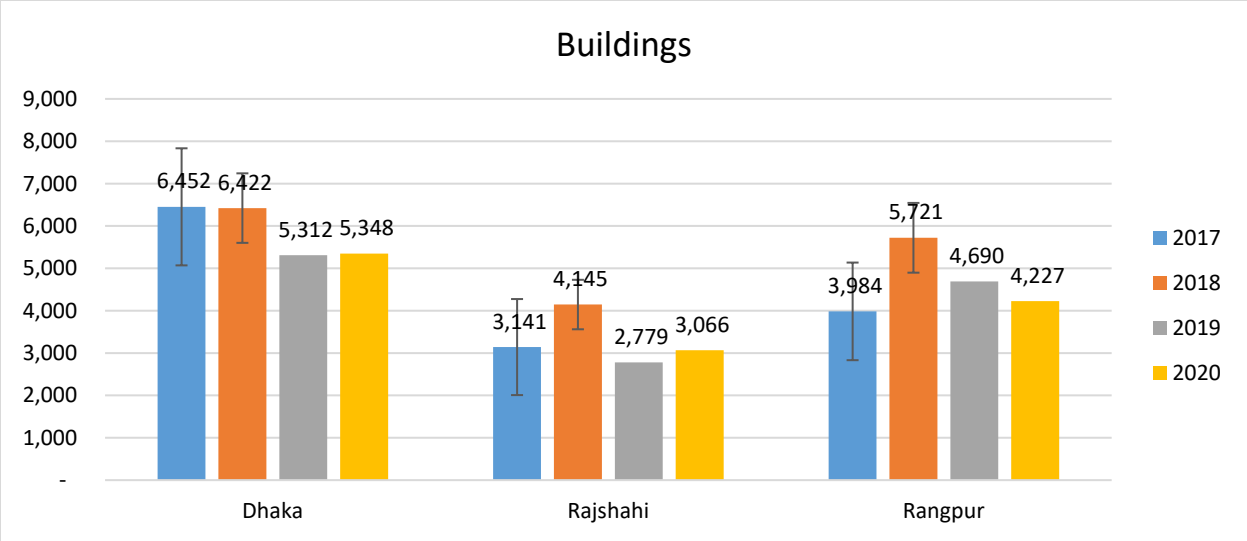

Figure 49 – Selection of optimal parameters. Variances of building classification for each year and province.

Figure 50 displays the building area using the optimal months and parameters as well as the average and standard deviation. The best model was the Random Forest algorithm with four and eight decision trees respectively for the Boro and Aman rice seasons.

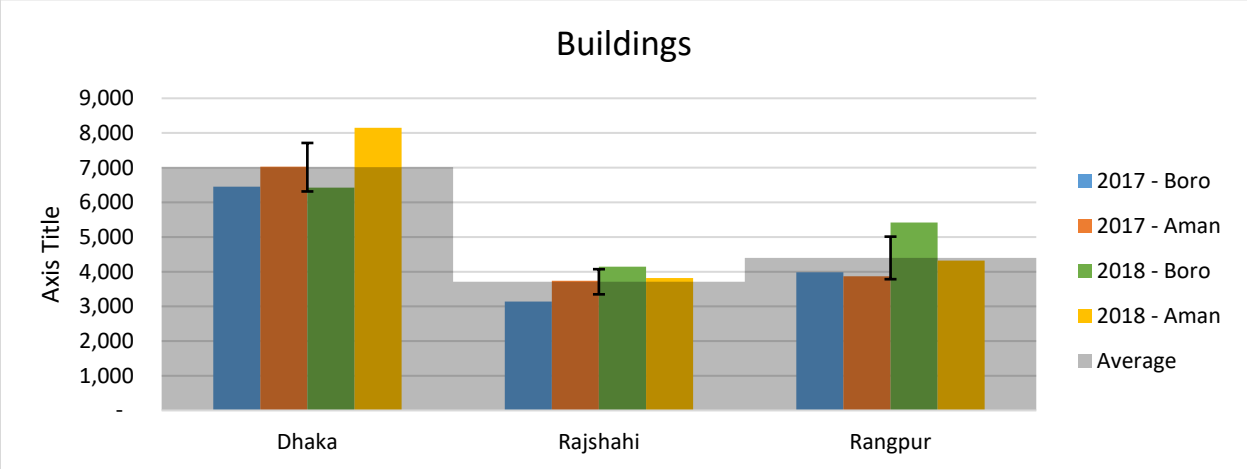

Figure 50 – Variance in building area estimates over the years.

The validation of the results consists of calculating the total area at the province level, and when possible, at a smaller sub-division for each rice season and comparing with official statistics. For comparison purposes, we conducted similar calculations using the Dynamic Land Cover map (CGLS-LC100), a new product in the portfolio of the Copernicus Global Land Service (CGLS). These Land Cover maps are provided at lower spatial resolution (100 meters) but for the period 2015-2019 and globally.

Results

We ran the analysis of Bangladesh, Indian, and Pakistan using the previously described model for several parameters. As previously mentioned, it is nearly impossible to conduct a systematic validation of the classification due to a lack of observed data, so we conducted a random visual inspection for improving the model by trial and error. The Random Forrest algorithm was the one exhibiting the best performance and the less variance when changing the training dataset, and the following results are for the Random Forest Algorithm with a hundred decision trees over four months of data for the Aman rice and over two months of data for the Boro rice. More results on the classification, validation steps, and misclassification are provided in Appendix 4.4.

Figure 51 summarizes the calculated rice area by growing season (Aman and Boro only, Aus was taken out) as well as the official statistics from Bangladesh and the official statistics from the FAO. The blue and orange indicate the growing season considered, the three black dash lines from 2015 to 2018 correspond to the harvested areas for the three growing seasons obtained from the Bengali Bureau of Statistics and the red squares correspond to the total area obtained from the FAOSTAT.

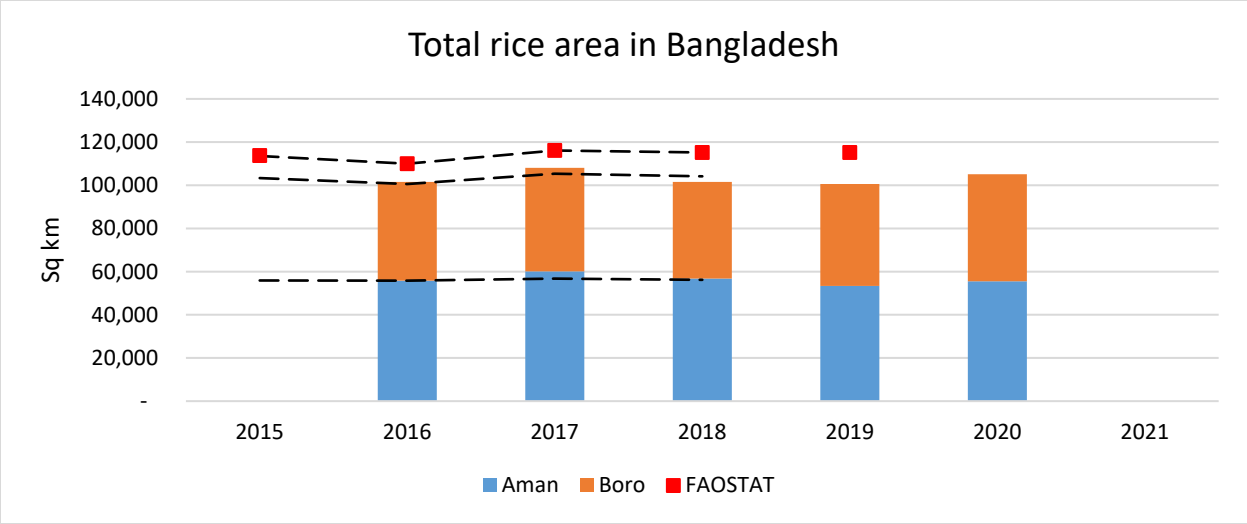

Figure 51 – Harvested area by year and season in Bangladesh. The black dashed line represents the Bureau of Statistics' data, and the red squares indicate the FAOSTAT data.

Figure 52 and Figure 53 summarize the finding for the Aman and Boro rice at the province level for each year. Overall, our results indicate a positive trend in harvested area in the Sylhet province for both growing seasons considered, a negative trend in Aman rice in the Barisal province, and significant annual variations without trends in the other provinces. Furthermore, we detected an increase in Boro rice in FY 2017/18 compared to FY 2016/17 (+5% harvest area), mainly in the provinces of Dhaka (+6.3%) and Sylhet (+61% harvested area).

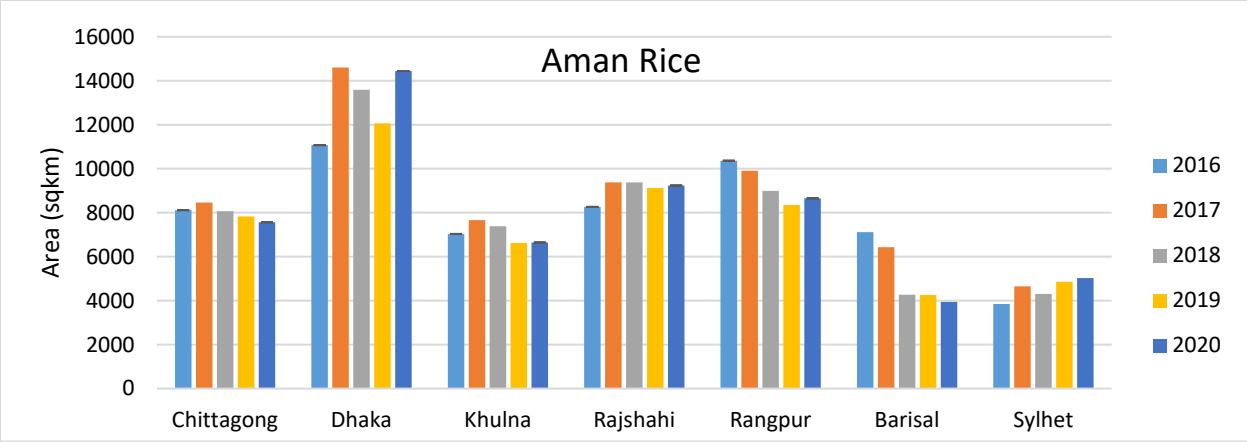

Figure 52 – Harvested areas of Aman rice by year and province.

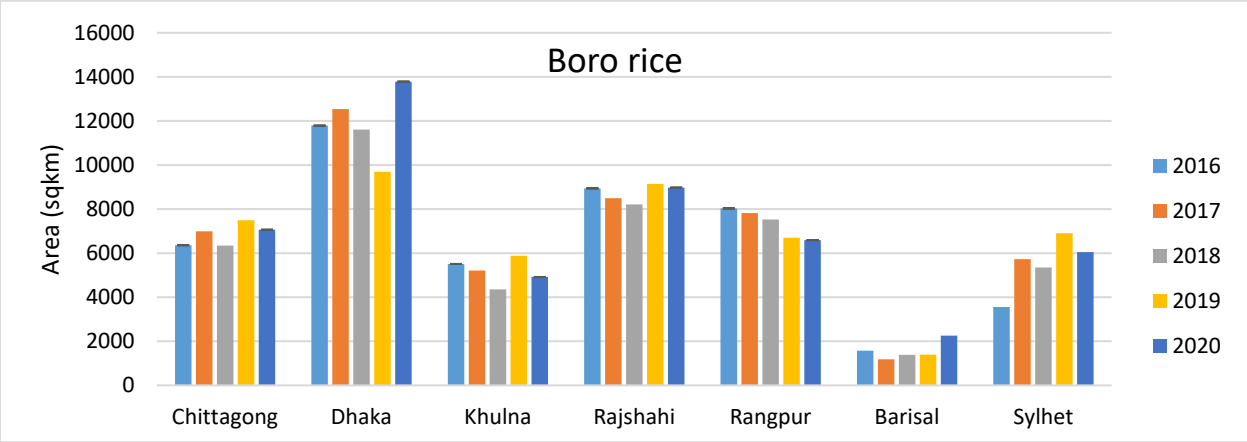

Figure 53 – Harvested areas of Boro rice by year and province.

Using the five years of harvested rice area, we quantified the percentage change over the five years defined as the ratio of the two most recent years to the two first years available at the district level. Trends are summarized in Figure 54 which reveals a strong decrease in the Dhaka district of the Dhaka province, in the coastal districts of the Barisal province, in one district of the Chittagong and Rajshahi provinces. Strong positive trends are also detected in all the districts of the Sylhet province.

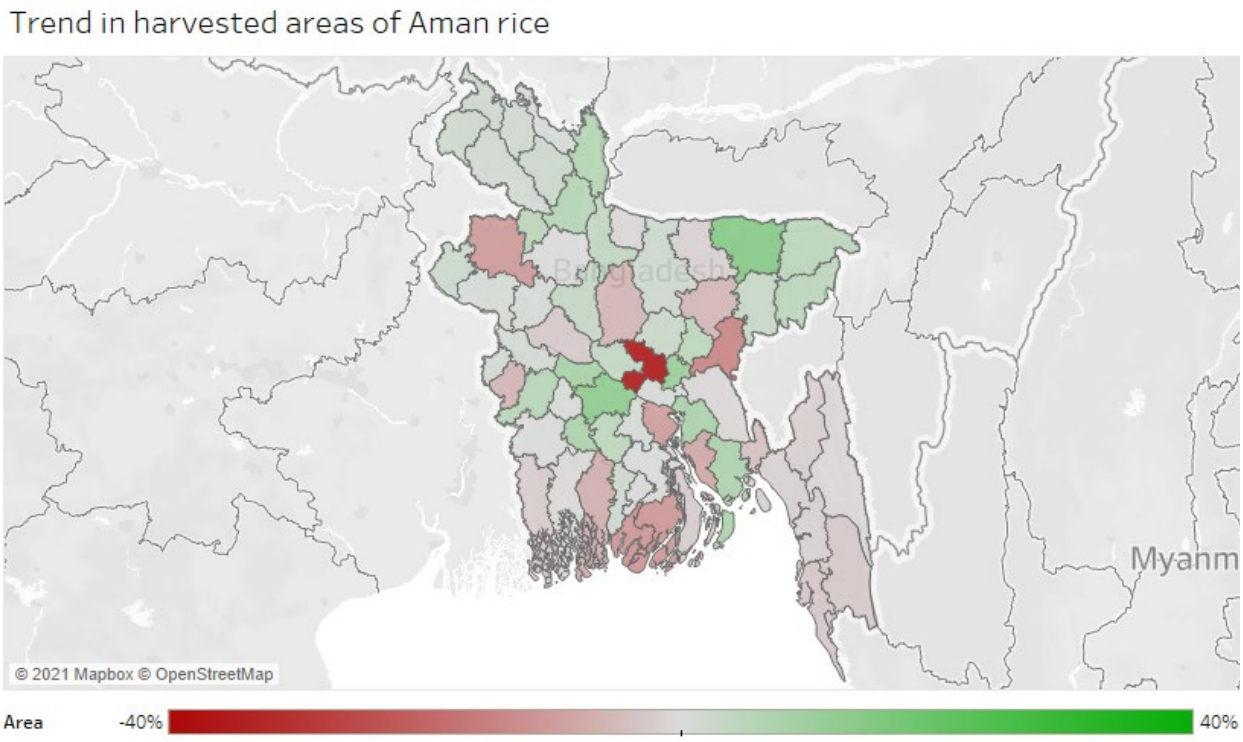

Figure 54 – Trends in harvested areas at the district level using the detected area of Aman rice.

Discussion

Using the validation dataset created from the annual harvested reports, we validated our findings of a positive trend in rice area in the province of Sylhet, with +23% for the Aman rice and + 48% for the Boro rice within the past five years (+10% and +33% for the Aman and Boro rice respectively from 2016 to 2018) and of a higher harvested area in Boro rice for FY 2017/18 compared to FY 2016/17 (+7% overall, +8% for the Dhaka province and +70% for the Sylhet province). The substantial drop in Aman rice detected in the coastal province of Barisal was also corroborated by official statistics.

We also compared our results with the Dynamic Land Cover from COPERNICUS (Buchhorn, et al., 2020) as shown in Figure 55. The DLC was imported to Google Earth Engine, and rice lands were extracted at division scales for analysis. The classification proposed in this chapter is differentiating between the Boro and Aman rice, therefore has more data points than the DLC, and exhibits a much higher correlation and lower error.

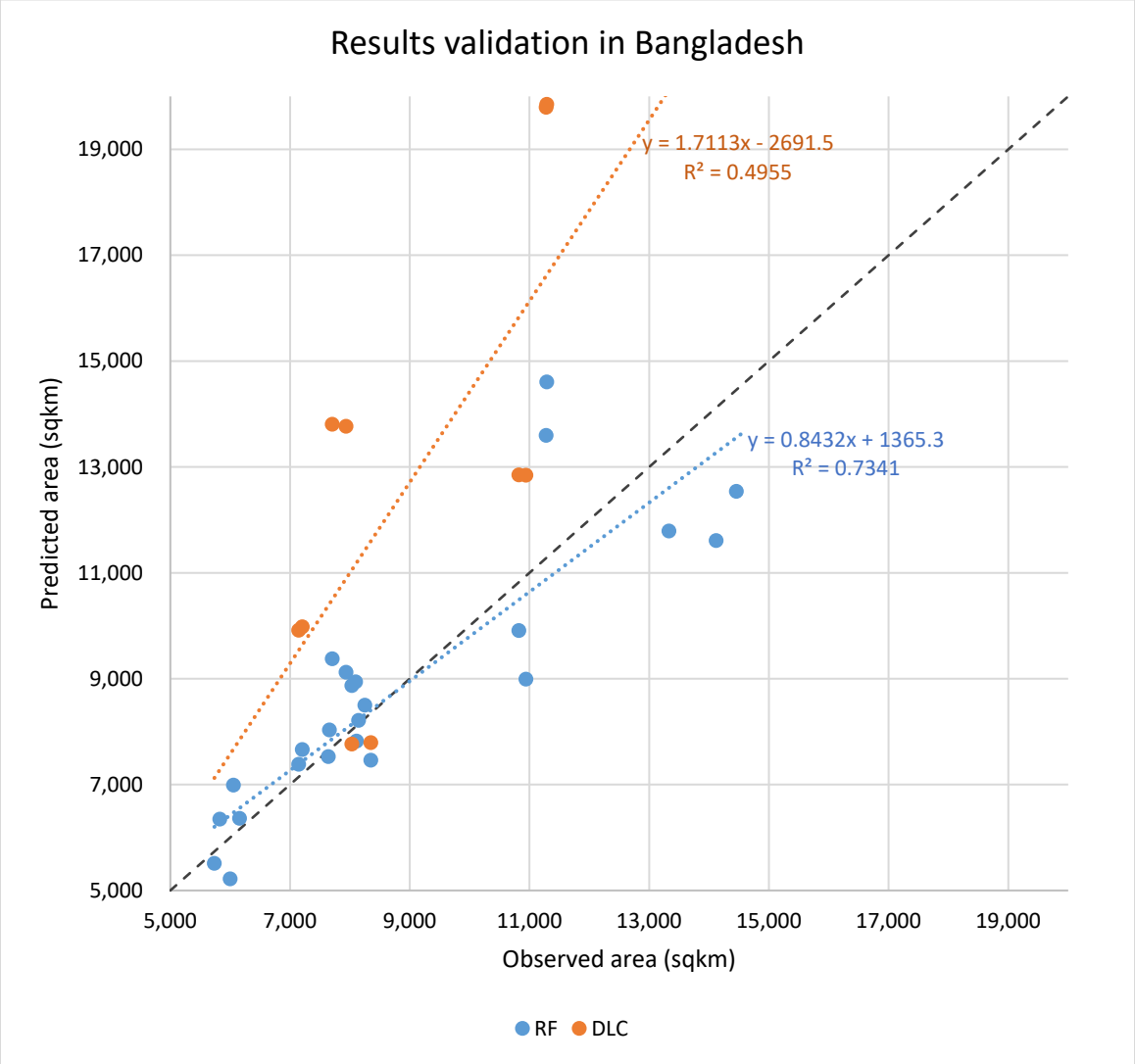

Figure 55 – Validation of the classification results and the Dynamic Land Cover to the ground observations. The blue color represents the classification proposed here detailed

for Boro and Aman rice and the orange colors represent the DLC from Copernicus incorporating all growing seasons.

Several drivers can explain the overall trends and year-to-year variations. In the coastal provinces, challenging growing conditions such as waterlogged soils and high salinity are preventing the use of High Yield Varieties (HYV) or the use of earlier maturing varieties of Aman rice. More information on rice yields can be found in Appendix 4.4. Yields are therefore lower than the national average, making rice paddies less profitable. Moreover, crops harvested are often time damaged or partially destroyed by pre-monsoon rains or cyclones or burnt by the sea-water intrusion. In 2018, losses were reportedly due to the spread of saline water up to 150km upstream in the province (Bhattacharya, Saha, Mondal, Bhandari, & Humphreys, 2019). These frequent destructions of rice paddies in the coastal Barisal province are adding pressure to already impoverished farmers and have been attributed as the main cause of the large migration from the Barisal province to the capital city Dhaka district (Dasgupta, S.; Hossain, M.; Huq, M.; Wheeler, D., 2014), and the resulting conversion of rice crop lands to urban and sub-urban areas. Biotechnological improvements and higher demands are also explaining the increasing trends in the Eastern Sylhet province. For instance, low concentrations of arsenic in the groundwater of the Sylhet province have been reported resulting in healthier and higher quality grains (Rmalli, Jenkins, Watts, & Haris, 2012). Arsenic is indeed a major groundwater contaminant in Bangladesh with concentrations always higher than the WHO's recommendations (Hossain, 2006), (Neumann, et al., 2009). Genes in local species of rice have also recently been

isolated and reported to increase the plant's tolerance to phosphorus-deficient soils (Gamuyao, et al., 2012).

Overall, our results are corroborated with official statistics at the district level, and several potential drivers of trends over these five years have been identified. Among them, soil salinity is one primary driver affecting yield (low productivity, fewer rice varieties adapted to the soil salinity) and harvested area (delayed maturing and harvesting, making rice paddies exposed to storms). The main mechanisms affecting soil salinity are tidal flooding during the wet season, critical to the Aman rice (June to October), and seawater intrusion during the dry season, critical to the Boro rice (November to May). Precipitations, especially lower than usual precipitations during drought or weak monsoon years, will also amplify the increase in soil salinity (Haque, 2006). Using ground observations and models, the median soil salinity in the coastal areas is expected to increase by 39% by 2050 (Dasgupta, Hossain, Huq, & Wheeler, 2014). As sea levels are rising and precipitation patterns are expected to change, there are growing uncertainties on the impact of salinity on food production in coastal areas.

Flood is also a major cause of loss of harvested area in Aman rice, especially flooding happening during the months of April to October, and flood impacts on rice harvest are well documented in the literature (Kanti Paul & Rasid, 1993), (Banerjee, 2010), (Rahman, Kang, Naghabhatla, & Macnee, 2017) and heavy rains and flash floods are often documented to negatively affect the Aman rice production (USDA, Foreign Agricultural Service).

Water and heat stresses are also major causes of loss of harvested areas, especially during the growing phases (January-April for Boro rice). For instance, the loss of rice production in the Dhaka region in 2016 was most likely attributed to high temperatures and low precipitation.

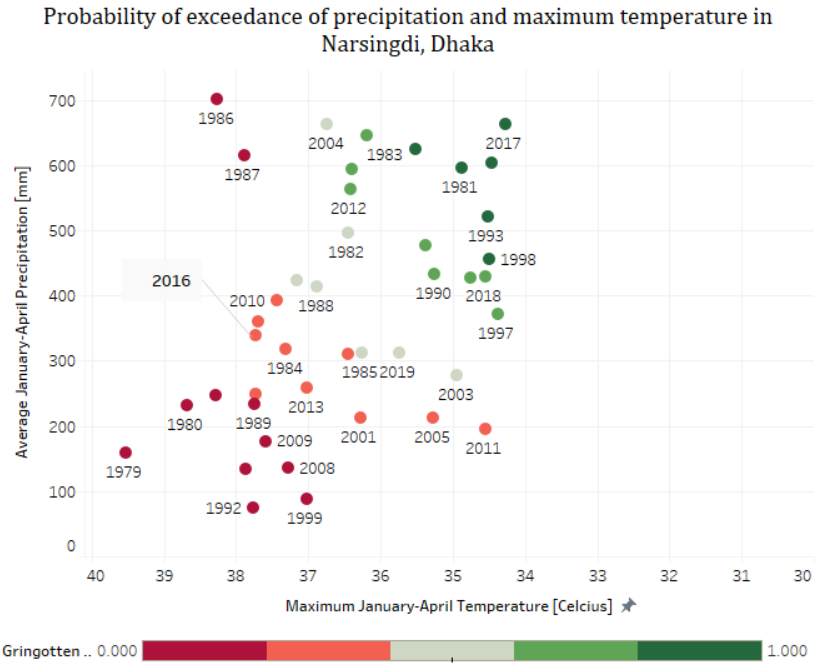

Figure 56 – Bivariate analysis of precipitation and maximum temperature. The colors indicate the probability of exceedance using the empirical copula defined by the Gringorten plotting position.

Figure 56 represents the probability of exceedance of low precipitation and high temperature. The temperature is the monthly maximum temperature from January to April, and the precipitation is the average amount during the same months. Data were extracted from the ERA5-Land is a reanalysis dataset using Google Earth Engine, the empirical bivariate distribution was defined using the Gringorten plotting position. The

calculated probability of exceedance in the district of Narsingdi in Dhaka seems to confirm the low harvested area in 2016 (relatively hot and dry) and the high values in 2017 and 2018.

Conclusion

This chapter demonstrated that past decreases in rice production in Asia were caused by both yield losses and reductions in harvested areas. We presented a machine learning pixel-based model to detect and quantify rice paddies during the winter and summer seasons (Boro and Aman rice) in Bangladesh using solely data remotely sensed by satellites. While there are abundant models mapping rice paddies using Sentinel-1 SAR or vegetation indices, our approach is more unique as it requires a minimum amount of training data, focuses on both Boro and Aman rice, and was optimized and validated using high-resolution ground observation of harvested area from 2017 to 2019 at district and province level in Bangladesh.

Our results show that pixel-based models using random forest classifiers can be used to accurately map rice paddies. We also showed that some provinces in Bangladesh recently significantly increased their cultivated rice areas for both growing seasons, while others experienced dramatic annual variations or steady negative trends. While attributing a cause to each variation is not trivial, higher demand and better access to new irrigation systems or more productive rice varieties are typically leading positive trends while diversification and accumulation of crop failures to negative trends. Flash floods, heatwaves, and saltwater intrusions are, on the other hand, responsible for annual variations.

We used high-resolution statistics from the Bureau of Statistics in Bangladesh at the district level (using all 64 districts of the eight provinces) for two growing rice seasons (Aman and Boro) for validation. While these results cannot be used for detecting long-term trends in harvested areas due to the short records, these results are essential to understand the dynamic changes in crop patterns, especially the dynamics between several districts within one region.

In this study, we used machine learning algorithms for crop mapping, which are powerful tools to do so, but we identified several weaknesses. First, high-resolution datasets are required to produced consistent and reliable crop maps or estimates of crop areas. With several growing seasons and diversity of planting strategies, high temporal resolutions are also required to detect and differentiate the three growing seasons. Moreover, diversification of crops, crops coupling, and change in planting strategies add challenges for ML algorithms to accurately map rice paddies, and more ground observations are required. Lastly, with climate change objectives, especially methane emissions from rice paddies, better access to irrigation systems, and crop diversification, models will pain to accurately map rice croplands, and higher spatiotemporal resolutions and ground observations will be needed to create, train, and validates future models.

CHAPTER 5

SUMMARY AND CONCLUSION

Extreme events can cause significant disruptions to the food systems and result in major financial and human losses. Droughts, heatwaves, storms, and floods dramatically impact crop productivity, farmers' revenue, and even consumers' nutrition security. Extreme events are typically seen as localized events with physical damages within a confined geographic area. Similarly, food security and resilience of a domestic food system are often treated as a domestic issue and rarely integrated other countries' food systems. Yet, food systems are increasingly globalized and increasingly dominated by a few major players characterized by intensive monocultures. Major food producers have consolidated their positions on the global food trade network, and their food products, main high yield varieties of cereals, have increasingly penetrated the food plates of other countries.

Consequently, with this growing concentration of cereals, several studies have assessed food security using different vulnerability indices. These indices provide valuable information on which countries are at the most risk of experiencing widespread malnutrition within their population. However, there is a strong focus of the scientific literature on yield, expressed in mass of food per area. Extreme events are known to negatively impact the productivity of the crops, effectively reducing the yield, but depending on the intensity and growing stage impacted, the quality could also be

negatively impacted, changing the nutritional values of the food produced or even reducing the harvest to zero.

This dissertation is an effort to develop a framework to track the impacts of one extreme event on other countries' food systems. The overarching goal is to better understand nonlocal food shortages and malnutrition by 1) modeling the impacts of one extreme event to other countries food supply using trade models and on nutrition supply; 2) using advanced machine learning tools and sun-induced fluorescence satellite imageries to predict cereals quality and protein content; and 3) using classifiers and high-resolution satellite imageries to map rice paddies and quantify trends and annual variations. In this dissertation, we are using only publicly and globally available data to be able to extend the framework to other countries. The machine learning and classifiers tools used a minimum amount of training data and were cross-validated for replicability in other regions. The framework built is separating the different components of the food supply chain so that this study can be generalized and extended to extreme events affecting other stages of this supply chain (e.g., hurricane on trade, flooding on storage quality, et cetera.).

In chapter 1 of this dissertation, the framework was applied to several extreme events in leading food producers, and the impacts of the food quantity supplied were calculated. We propagated a change in production through the global trade network and quantified the changes in the food supply in other countries. The proposed trade model is a network-based model calibrated using historical data for accounting for an individual country's reaction to a large production or a large importation deficit. Our results are

corroborated by our preliminary study on the countries the most at-risk using diversity indices, and we showed that extreme events in one major producer could dramatically impact food security in other countries.

In Chapter 2, we create a mass-balance algorithm based on the Food and Agriculture Organization – FAO – to track changes in nutritional values of the food produced, traded, and consumed. We create a dataset and quantify the impacts of a variation of the nutritional value of several food products. Our analysis showed that nonlocal variations of the nutrition values of even wheat only could dramatically impact malnutrition in several countries. In other words, if all wheat consumed and not locally produced had a 15% lower protein content than the standard values, several countries which are often time presumed to meet their dietary energy supply would be under their calory requirements. We also showed that major food exporters could drastically impact other countries' malnutrition. For instance, a decrease of 15% of the wheat protein content in the USA only would significantly impact calory and protein supply in other parts of the world. The dataset created offers a unique opportunity to test a country's risk of malnutrition due to nonlocal changes in crops' quality.

Chapter 3 proposed an algorithm to forecast the wheat protein content with one month lead time in the United States. We developed a database of ground observation based on harvest reports of Hard Red Winter wheat in the USA from the past two decades. Our algorithm was trained and tested over a decade of data to cover several growing

conditions (droughts, heat stress, normal, wet…) and provide more robust predictions. We used only precipitation and Sun-Induced Chlorophyll Fluorescence satellite imagery to predict the wheat protein content. Our results indicate that the best results are obtained one to two months before the harvest. Using Sun-Induced Fluorescence is promising to monitor the photosynthesis activity of the crop during the growing phases corresponding to the remobilization of the Nitrogen N from the vegetative parts of the wheat to the grain. Although the algorithms have been tested only in the United States, where crop fields are large enough relative to the satellite product resolutions, this methodology can be deployed to other countries using either high-resolution data from Unmanned Aerial Vehicles (UAV) or the future FLEX Mission from the European Space Agency.

In the last chapter 4, we proposed a pixel-based classifier algorithm to detect rice paddies using only satellite imageries from Sentinel-1 at 10 meters resolutions. While many studies are using SAR for rice mapping, our approach is unique as we looked at annual variation and trends over five years in Aman and Boro rice in Bangladesh. Moreover, the algorithm doesn't require an enormous amount of training data, and the trained algorithm can be deployed in regions outside of the training area with similar climatic conditions. We built a database of harvested areas of Aman and Boro rice at the district level in Bangladesh to validate our results. In this chapter, we tried to attribute a cause to several trends and annual variations. We concluded that saltwater intrusions were the leading causes of loss of harvested areas in the coastal districts (burning crops) but also in the capital city's district (large scale migrations of coastal farmers to the capital city and

consequently conversion of rice paddies to buildings). The migration of ruined farmers from the coast to the capital city illustrates this dissertation's objective to quantify nonlocal food production shocks and extreme events impacts beyond their physical boundaries. We also concluded that drought and heatwaves were responsible for a considerable loss in Boro rice while flash floods mainly affected the Aman rice.

In general, our framework can be generalized and extended to several types of rare events. 2020 has been a year of a couple of new and untraditional events impacting food security, such as labor shortage reducing agriculture productivity, cyberattacks affecting meat processing plants, ships wreckage affecting marine transport through the Suez Canal. There are many opportunities to investigate such events for scenario-based analysis and build probabilistic risk indices of malnutrition globally. Yet, there are several challenges to address as well and opportunities to consider compound or cascading events. For instance, seawater intrusion can be expected to increase due to sea-level rise and changes in precipitation patterns, resulting in lower yield, but also delayed harvesting, potentially exposing the crops to new hazards such as storms and high monsoon precipitation.

Furthermore, considering climate and environmental goals being constantly raised, and the growing population and the changing demographics, food systems will soon look different. Crop's diversification and especially crop coupling will make it harder for satellite imageries and machine learning algorithms to differentiate between several types of crops. High-resolution multi-sensors multi-temporal imageries will be needed to monitor crop production and quality.

APPENDIX

Appendix 0.1. Nutrition security

Figure 57 shows the importation dependency ratio for both calories and proteins.

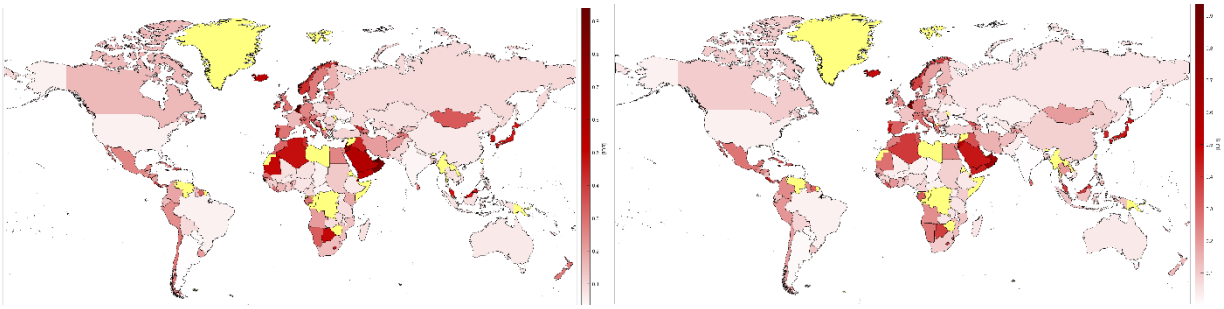

Figure 57 – Nutrition Importation Dependency Ratio average from 2000 to 2010. Left: For calories. Right: For proteins

Figure 58 shows the total export of corn by the USA and the port of departure. It reveals the critical roles of the Mississippi River in moving corn from the producing regions (US Midwest) to the port of departure (Louisiana) to the commercial partners (rest of the world). Roughly 80% of the exported US-corn is transiting by the Mississippi River.

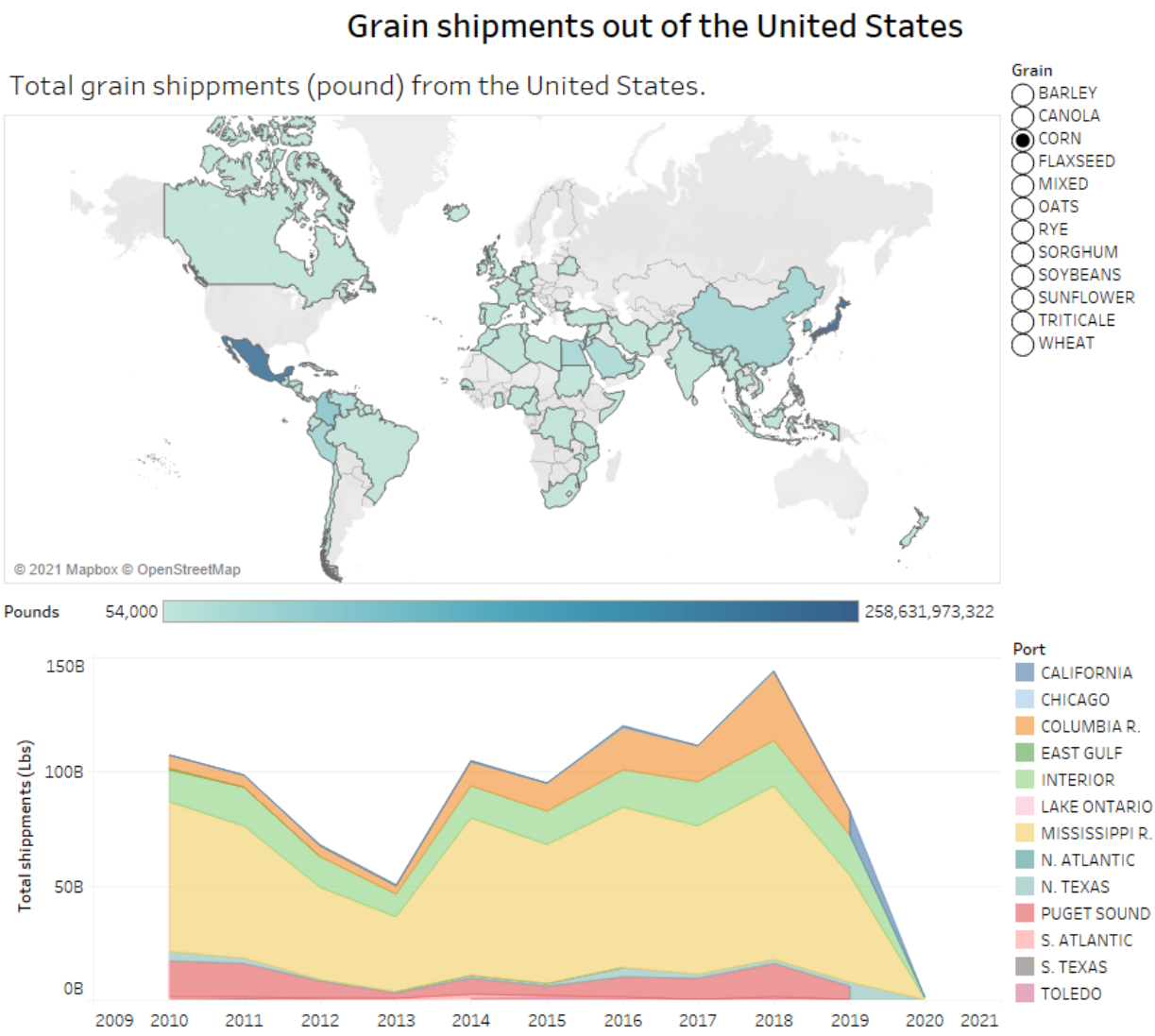

Figure 58 – US corn grain shipments by port and destination. Global export of US (top chart) and preferred port of departure for corn (bottom chart).

However, disruption of the traffic is not rare: hurricanes in the Gulf, flooding the Mississippi. Such disruption would be translated into a loss of grain for corn farmers if they must store the harvest on the unequipped field for too long, of increase of costs to sell the grain if alternative transportations are used (such as trucks or trains). The same analysis

for wheat can be found in Figure 59, in which case, the Columbia River is more than 50% of the total exportations.

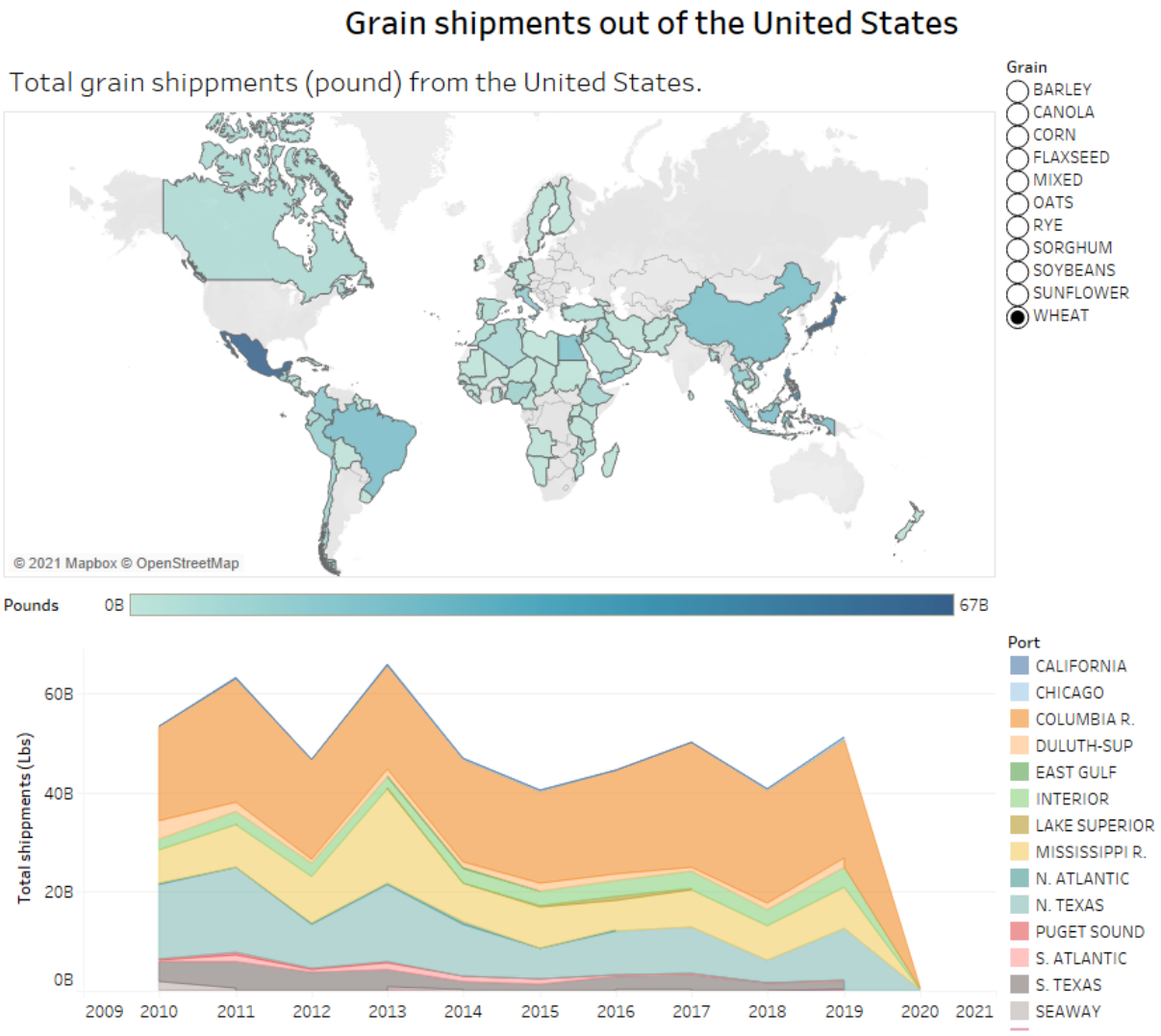

Figure 59 – US wheat grain shipments by port and destination. Global export of US (top chart) and preferred port of departure for wheat (bottom part).

Appendix 0.2. Ozone parameters

There are several surface ozone indices used in the literature for assessing the impacts on crop yields, and they are summarized in Table 6.

Table 6 – Most used ozone indicators for assessing crop damages.

Ozone index	**Definition**
AOT40 (nmol/mol.hr)	Sum of hourly daylight (>50 W/m2 PAR) ozone volume mixing ratio (VMR) above 40 nmol/mol.
SUM06 (nmol/mol.hr)	Sum of 24-houtlry ozone VMR above 60 nmol/mol.
W126 (nmol/mol.hr)	Weighted sum of 24-hourly ozone VMR.

The accumulated exposure over a threshold (AOT) is widely used in Europe with a threshold of 40ppb for crops (AOT40 Index) (Fuhrer, Skarby, & Ashmore, 1997). It is not clear if the limit should be 30 ppb or 40 ppb and it has even been found that the best results are AOT10 for wheat, AOT110 for grass-clover mixtures (Pleijel, et al., 1996). Moreover, AOT40 considers only daylight ozone concentration, and an important remaining question is whether the nighttime ozone concentration can contribute to damaging the plants. Another metric has been proposed (Karlsson, et al., 2007). Furthermore, the AOT40 is looking only at grain or seed yield rather than using photosynthetic rate, biochemical changes, or change in disease susceptibility. In the USA, SUM06 and W126 are favored. SUM06 considered only the concentration above 60ppb and then accumulates the total concentration. W126 uses a continuous rather than step-weighting function, with a sigmoidal distribution, i.e., weights of 0.03, 0.11, 0.30, 0.61 and 0.84 at ozone volume mixing ratios of 40, 50, 60, 70, and 80 ppb, respectively.

Appendix 0.3. Trends in cereal yields

Figure 60 reveals the trend of wheat's yields globally using 30-years of data obtained from FAOSTAT and using a Mann-Kendall test for trend at 95% confidence interval.

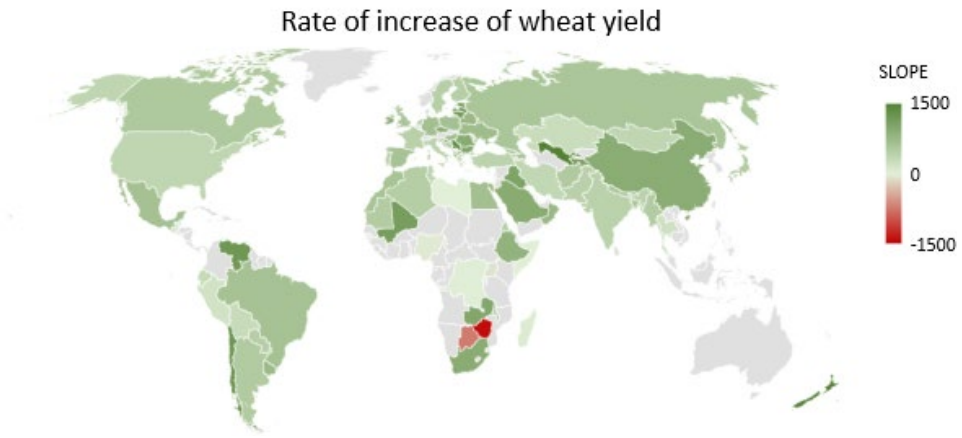

Figure 60 – Trends in wheat yield from 1990 to 2019. Values of 0 indicate no significant trend detected using a Mann-Kendall test with a 95% confidence.

Yield time series were then detrended, the departure from average computed, and ranked by magnitude (5%, 10%, or 20% departure from the average) as shown in Figure 61.

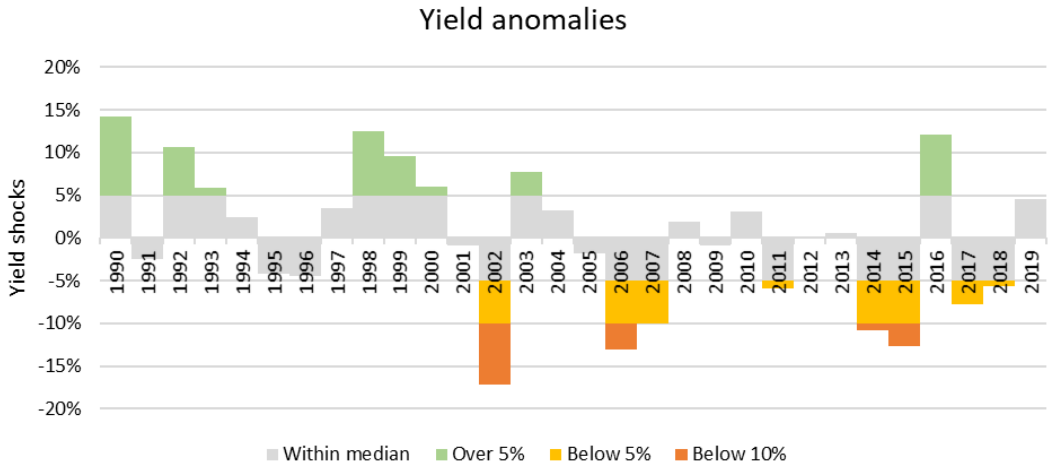

Figure 61 – Anomaly in wheat yield by magnitude. Data from 1990 to 2019.

Appendix 1.1. Additional results

We calculated several vulnerability indices to point out the countries the most at risk from nonlocal production shocks and compared them with our model's findings. Figure 62 represents the vulnerability index from the United States, average from 2000 to 2010. The Vulnerability Index (VI) of the country i is defined as:

$$VI(i) = \sum_{k} IDR_k^i \cdot HHI_k^i \cdot w_k^i$$

Where k is a crop, IDR the Importation Dependency Ratio previously defined, HHI the market concentration index, and w the relevance of the crop k to the total calory supply of the country i.

The vulnerability index indicates a country's capacity to stabilize its dietary energy supply in case of a global and uniform decrease of US supply.

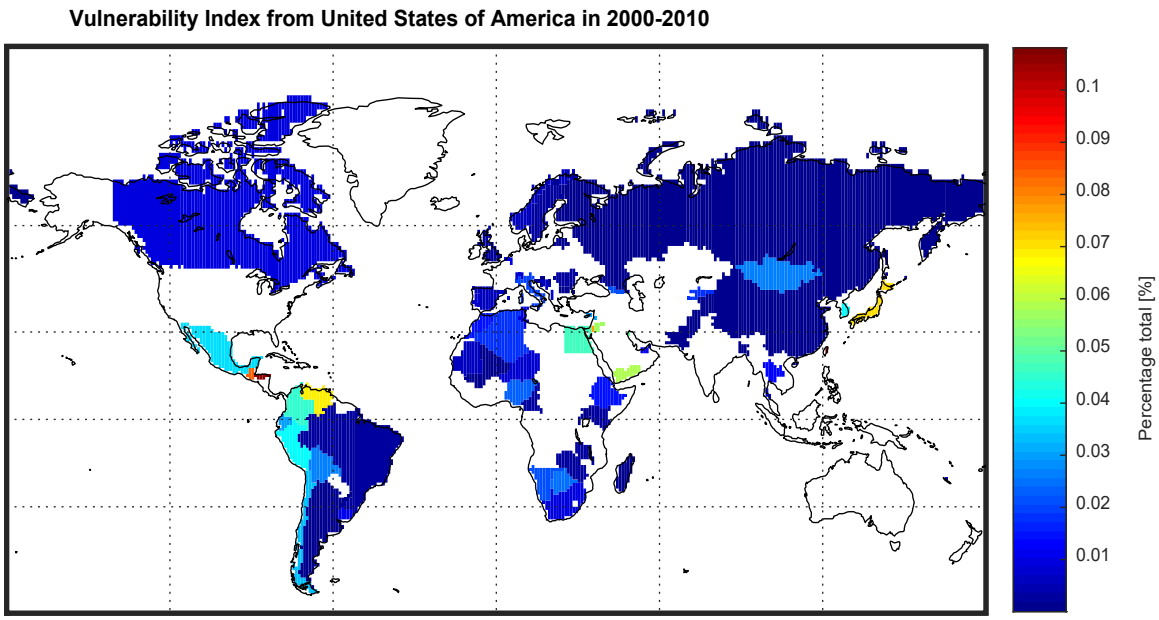

Figure 62 – Vulnerability Index from the US Wheat. A lower number indicates a higher change to stabilize calory supply in case of a global decrease of US exportation in the USA.

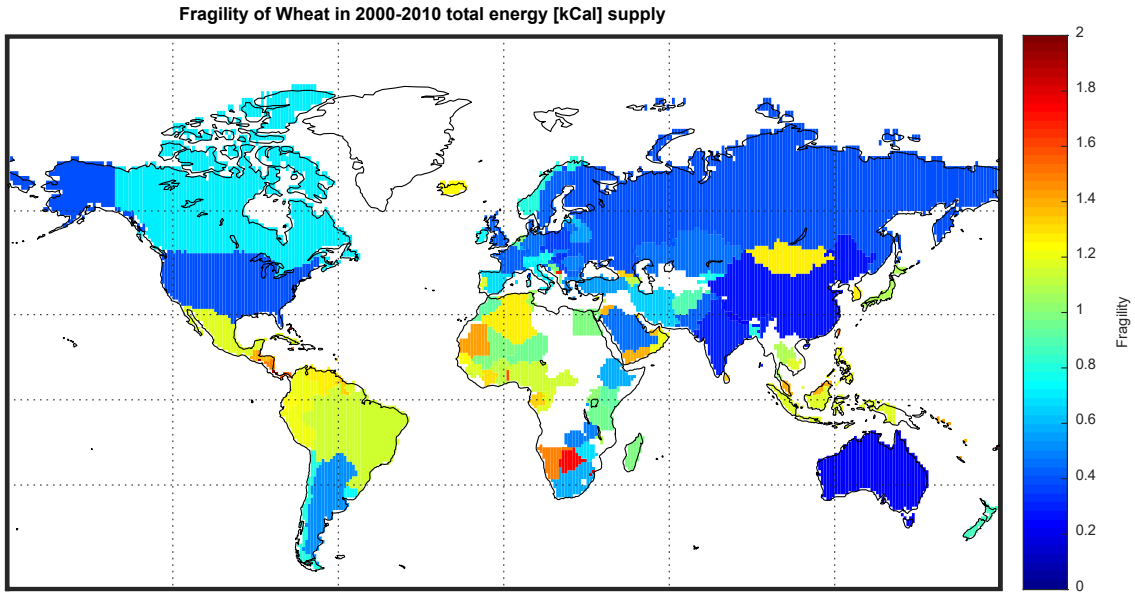

Figure 63 – Country's wheat fragility. A higher number indicates a higher risk of not being able to stabilize dietary energy supply in case of a global decrease in wheat exports.

Appendix 2.1. Model parameters

Figure 64 represents the food utilization ratio, defined as the ratio of the number of wheat products (wheat and flour) for human consumption by the quantity of wheat for other usages (feed, seed, processing) or wasted. The ratio is calculated from FAO's estimate of the lost products (waste) as well as seed, feed, and processing.

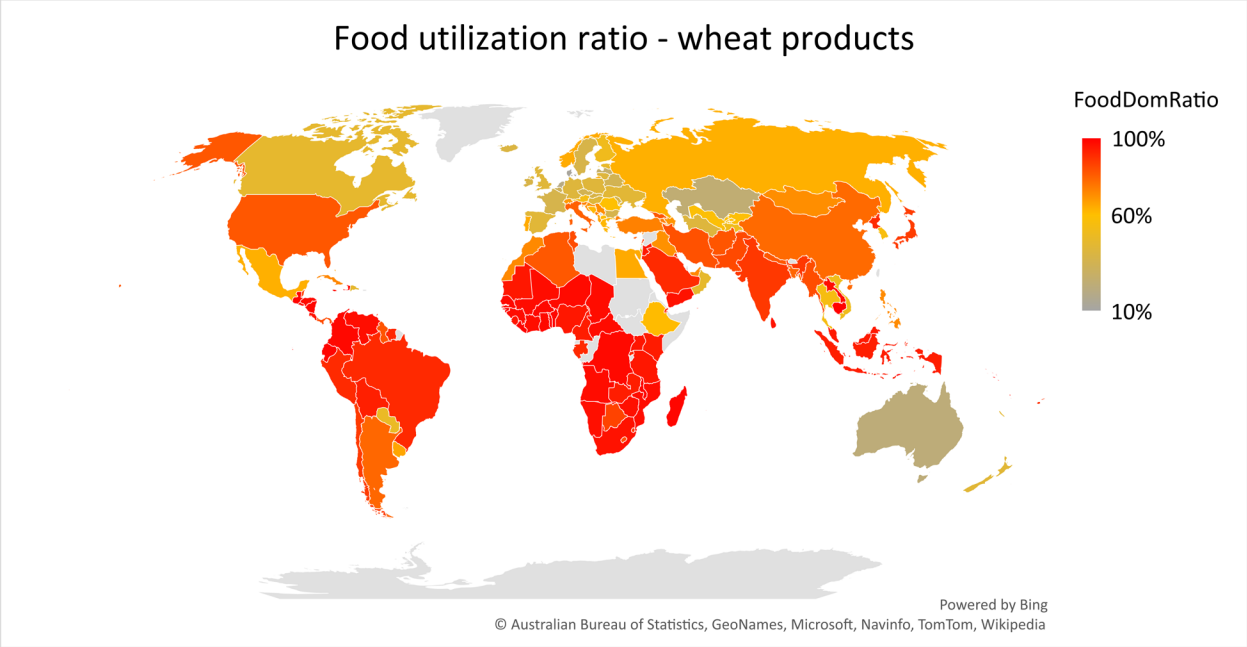

Figure 64 – Food utilization ratio of wheat in 2010, i.e., percentage of wheat used for food purposes.

We calculated the contribution of the national production, storage capacities, and trade flows to the domestic supply for food consumption. We define the coefficient α, β, and γ such as:

$$Food = \alpha \cdot Production + \beta \cdot \Delta Storage + \gamma \cdot \Delta Trade$$

Figure 65, Figure 66, and Figure 67 represent this coefficient for wheat products in 2010. A coefficient of 100% means that all the commodity was used for food. It provides further information on food security, for instance, how much of the domestic production is used for human consumption, how much of the storage, and how much of the import. Countries in red in Figure 67 are countries using all their wheat imports for human consumption, meaning that they will have few margins of maneuver to feed their population in case of a global decrease of traded wheat.

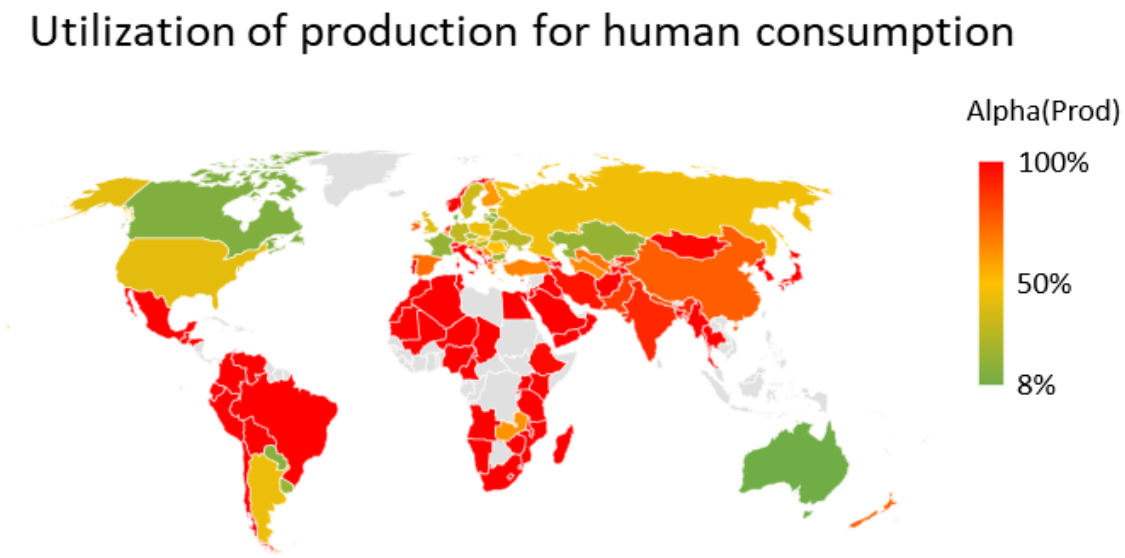

Figure 65 – Share of wheat's production for food supply: alpha coefficient.

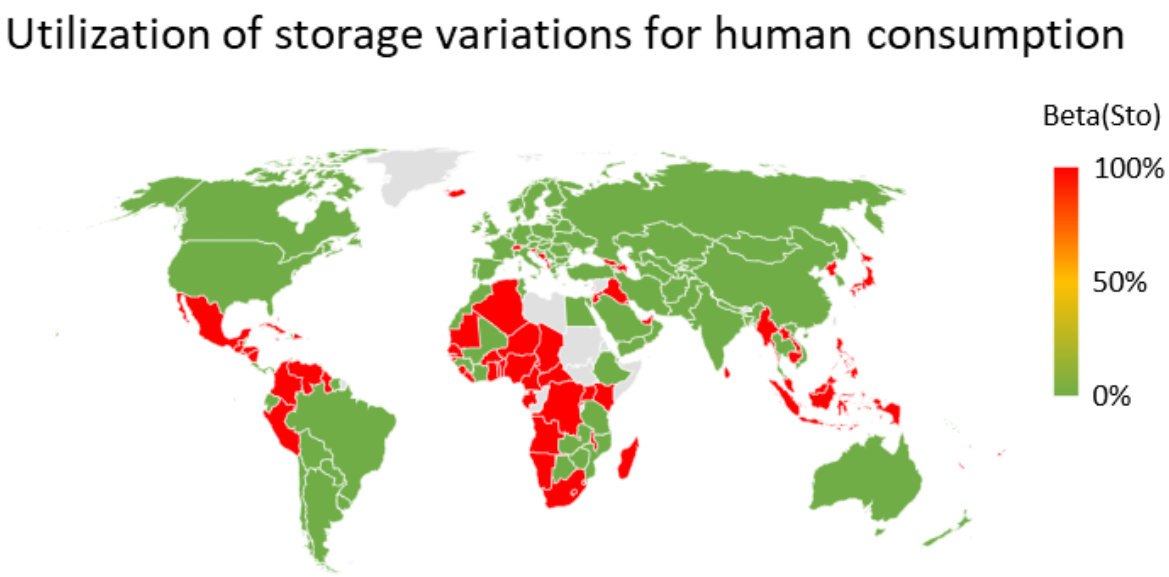

Figure 66 – Share of wheat's storage for food supply: beta coefficient.

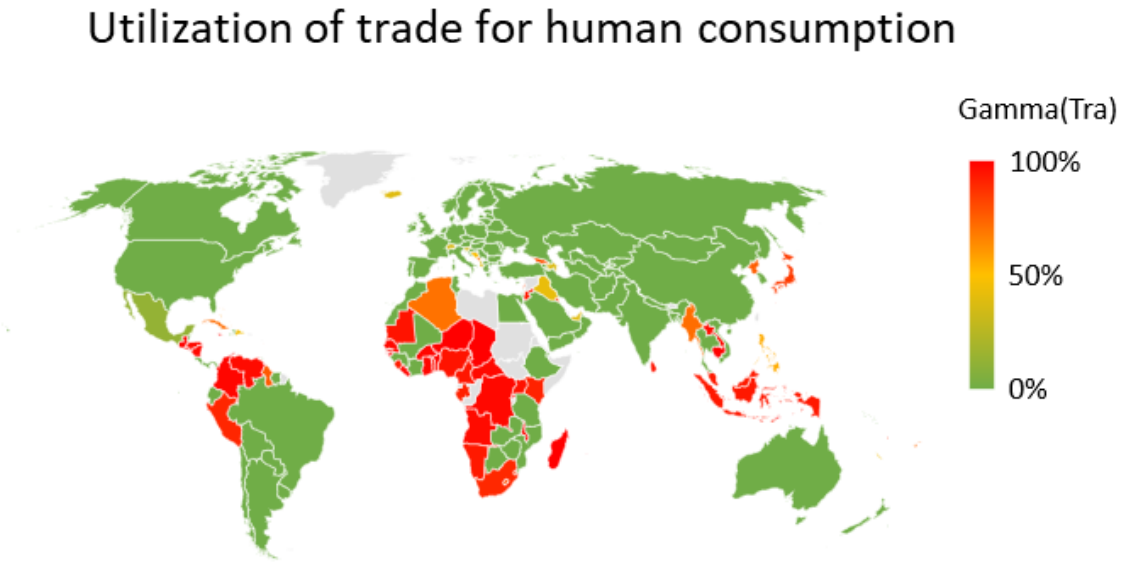

Figure 67 – Share of wheat's trade for food supply : gamma coefficient.

Appendix 2.2. Data availability

One major challenge to evaluate the nutritional value of the food supplied is the availability of an annual dataset at the country scale. Such data are only available in a few countries (Table 7).

Table 7 – Data availability and source for some of the most produced cereals.

Producing agency	**Commodity**	**Spatial Coverage / division**	**Time coverage**	**Element included**
Plains Grain Inc.	Wheat (Hard Red Winter)	United States / agricultural district	2009-2020	Protein
U.S. Grains council	Corn	United States / national	2015-2020	Protein, starch
U.S. Grains council	Sorghum	United States / national	2015-2020	Protein
US Wheat	Wheat (All class)	United States / national	2000-2020	Protein
Canadian Grain Commission	Wheat (All Class)	Canada / Province	2011-2020	Protein
France Agri Mer	Wheat (All Class)	France / Region	2000-2020	Protein

Appendix 3.1. Crop calendar trends.

The duration of some critical stages of the HRW wheat crops are typically around 30 days but have annual as well as geographical variability. The following figures (Figure 68, Figure 69, Figure 70) represent the median operation progress for HRW wheat over Kansas by week number. The three horizontal lines correspond to 50%, 75%, and 95% progress. Using weekly crop progress from surveys from 1982 to 2002 obtained from the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), we also included the 75th and 95th percentile. Besides 2012, which was exceptionally early, annual variability is typically one to two weeks.

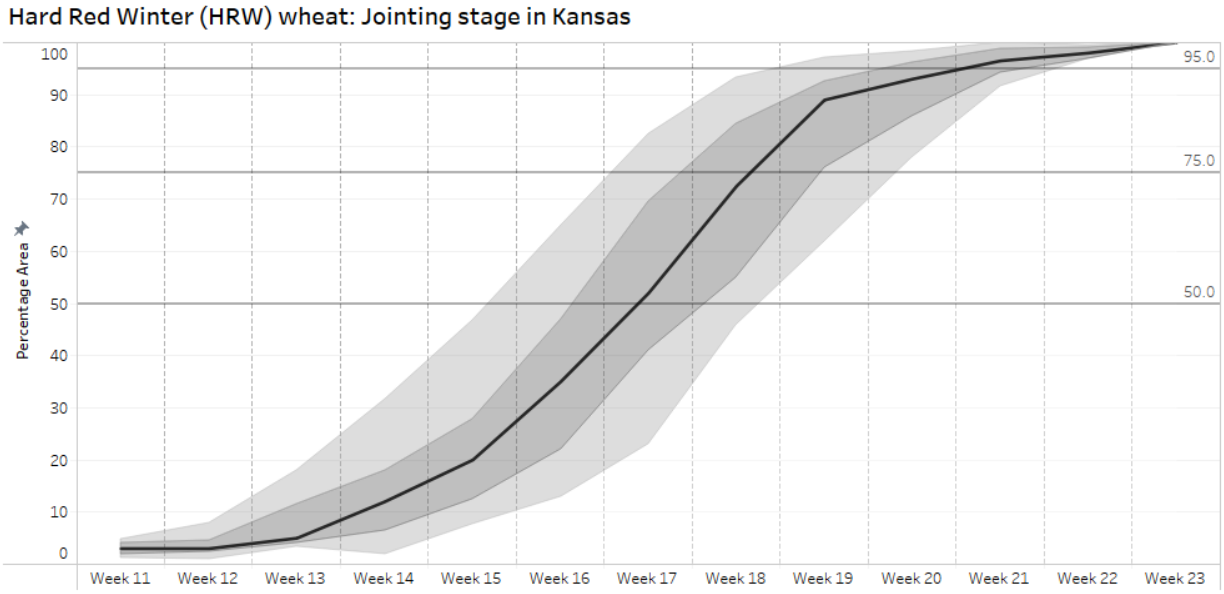

Figure 68 – Progress of the jointing stage for the Hard Red Winter (HRW) wheat in Kansas.

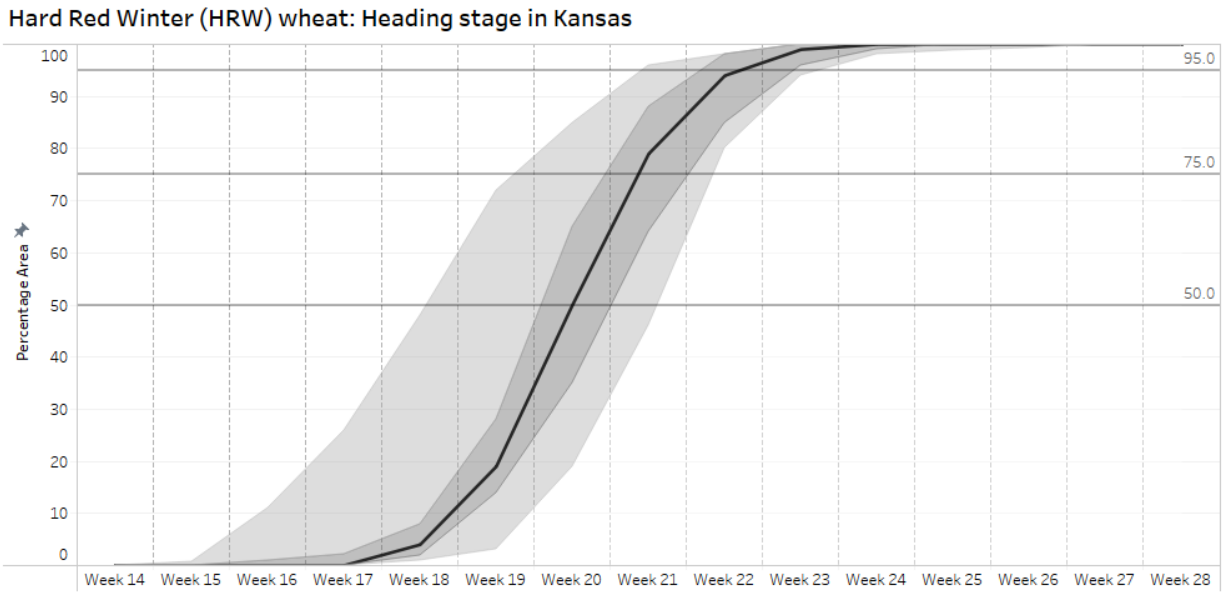

Figure 69 – Progress of the heading stage for the Hard Red Winter (HRW) wheat in Kansas.

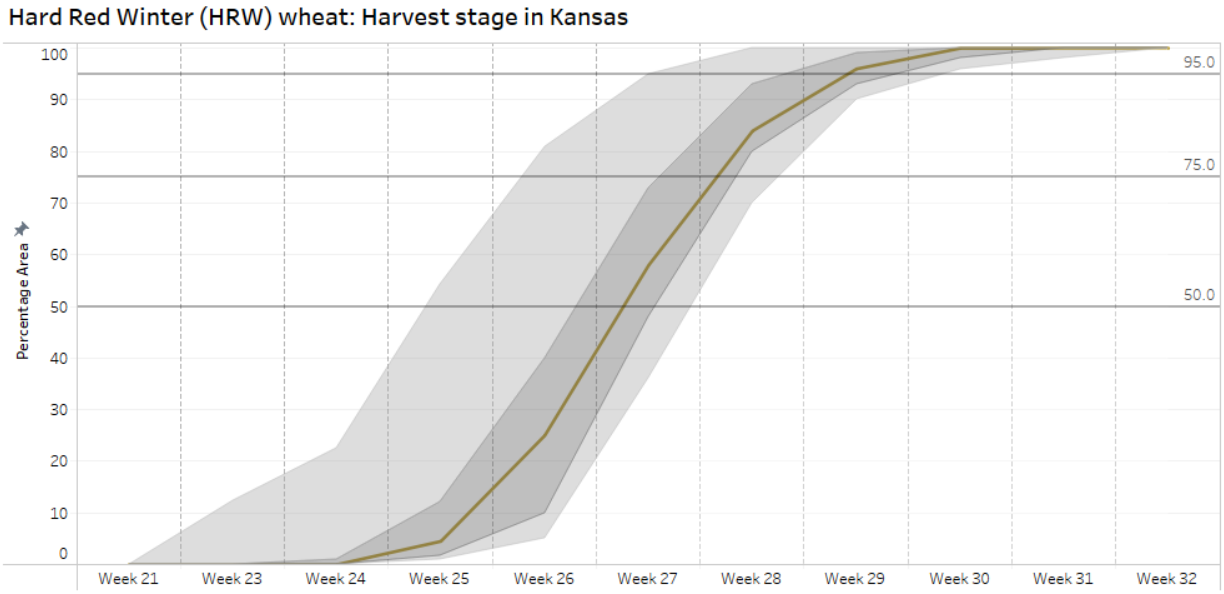

Figure 70 – Progress of the harvest stage for the Hard Red Winter (HRW) wheat in Kansas.

In addition to annual variability, we tested the data for long-term trends. Global warming could indeed affect the regional temperature and precipitation patterns, and therefore the date of the optimal growing season. We detected a positive trend in the planting progress

of the HRW wheat of about 3 weeks from 1981 to 2020, which corresponds to roughly one week every 10 years, however, these long term trends are relatively small compared to years from years variations or state to state difference (Figure 71 and Figure 72). The three lines represent the week number during which 25%, 50%, and 75% of the crop operation are reached.

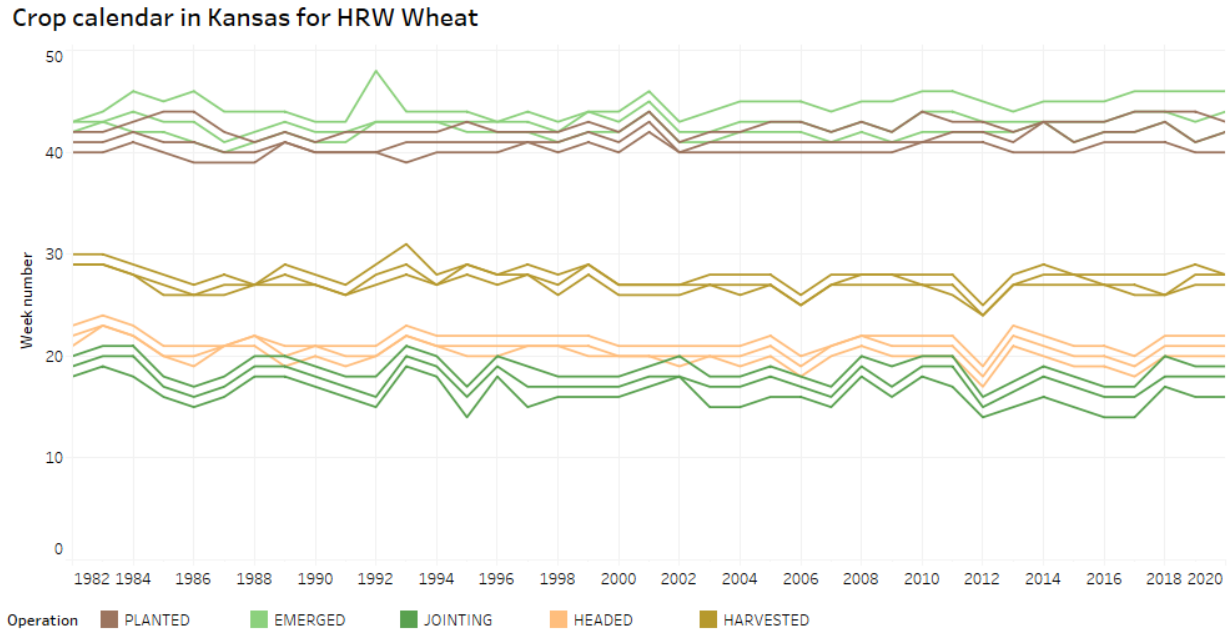

Figure 71 – Weekly progress of crop operations in Kansas. Week number of 25%, 50%, and 75% progress of selected Hard Red Winter Wheat operations.

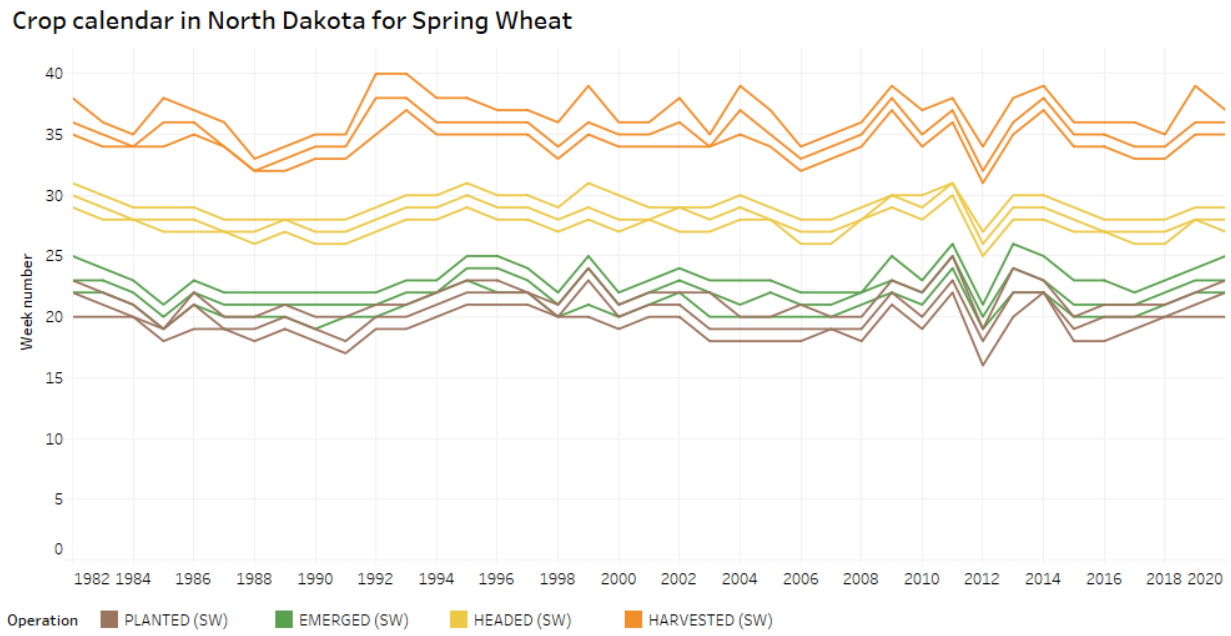

Figure 72 – Weekly progress of crop operations in North Dakota. Week number of 25%, 50%, and 75% progress of selected Spring Wheat operation.

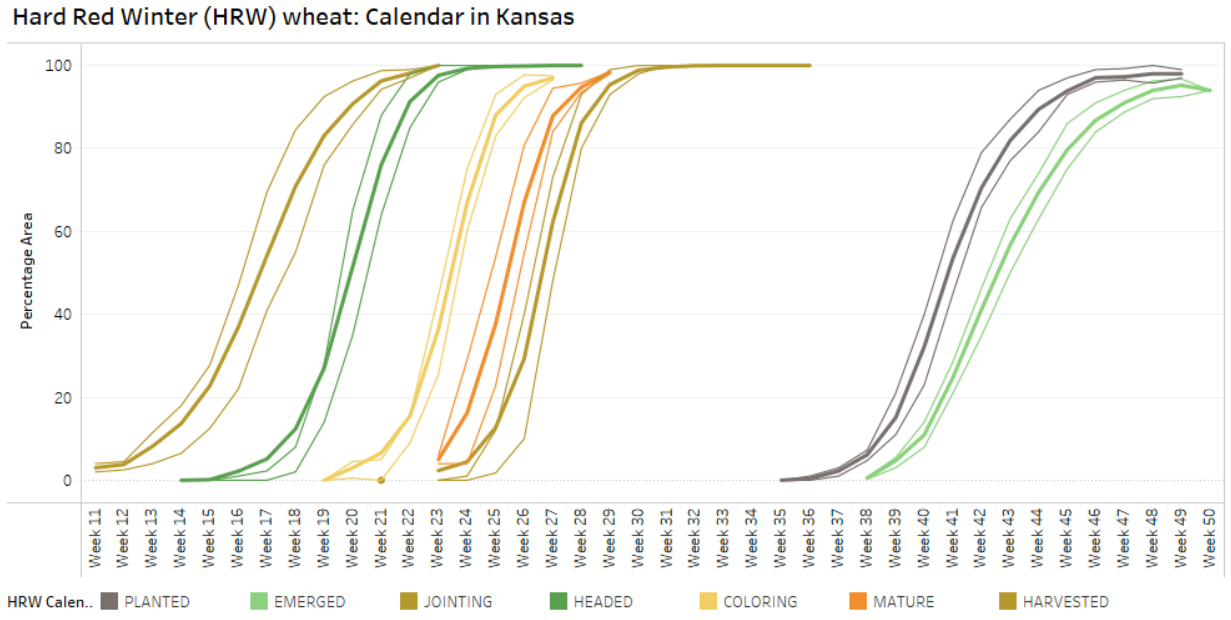

Figure 73 – Weekly progress of the main operations on Hard red Winter Wheat in Kansas.

Figure 73 displays the average and 25th, 75th percentile weekly progress of the main operations on the Hard Red Winter Wheat grown in Kansas.

Appendix 3.2. Model definitions

For each model, we chose to use a support vector machine-based regression model (Support Vector Regression or SVR). Support Vector Machines (SVM) are well-known classification algorithms, adapted for regression. Unless regular linear regression models, which aim to minimize the sum of squared errors (with or without penalty terms), SVRs are aiming to fit the error term within a certain threshold, which allows for more flexibility. It is an extension of a supervised classification algorithm, trying to classify the predicted values between the one inside or outside of the error boundaries and to minimize the norm of the vector. The following figures in Figure 74 explain the functions of the error term. The top plots are using an excessively large error parameter (value of 0.5), resulting in most of the values being fitted within the support vector (within 0.5 from the vector) and poor performance for the testing. The bottom plots are using excessive small error parameters (value of 0.001) resulting in few values attracted to the support vector and very spread values on the testing phase. The middle plots are using a moderate error factor (values of 0.1 corresponding to a whole range of 0.2% protein) and exhibit a more consistent performance between the training and the testing models. Since we are predicting protein content, the error parameter can be estimated from the data we have. A value of 0.5 corresponds to a range of 1g/100g protein content, which is too high. We estimate that a value of 0.2g/100g protein content was acceptable.

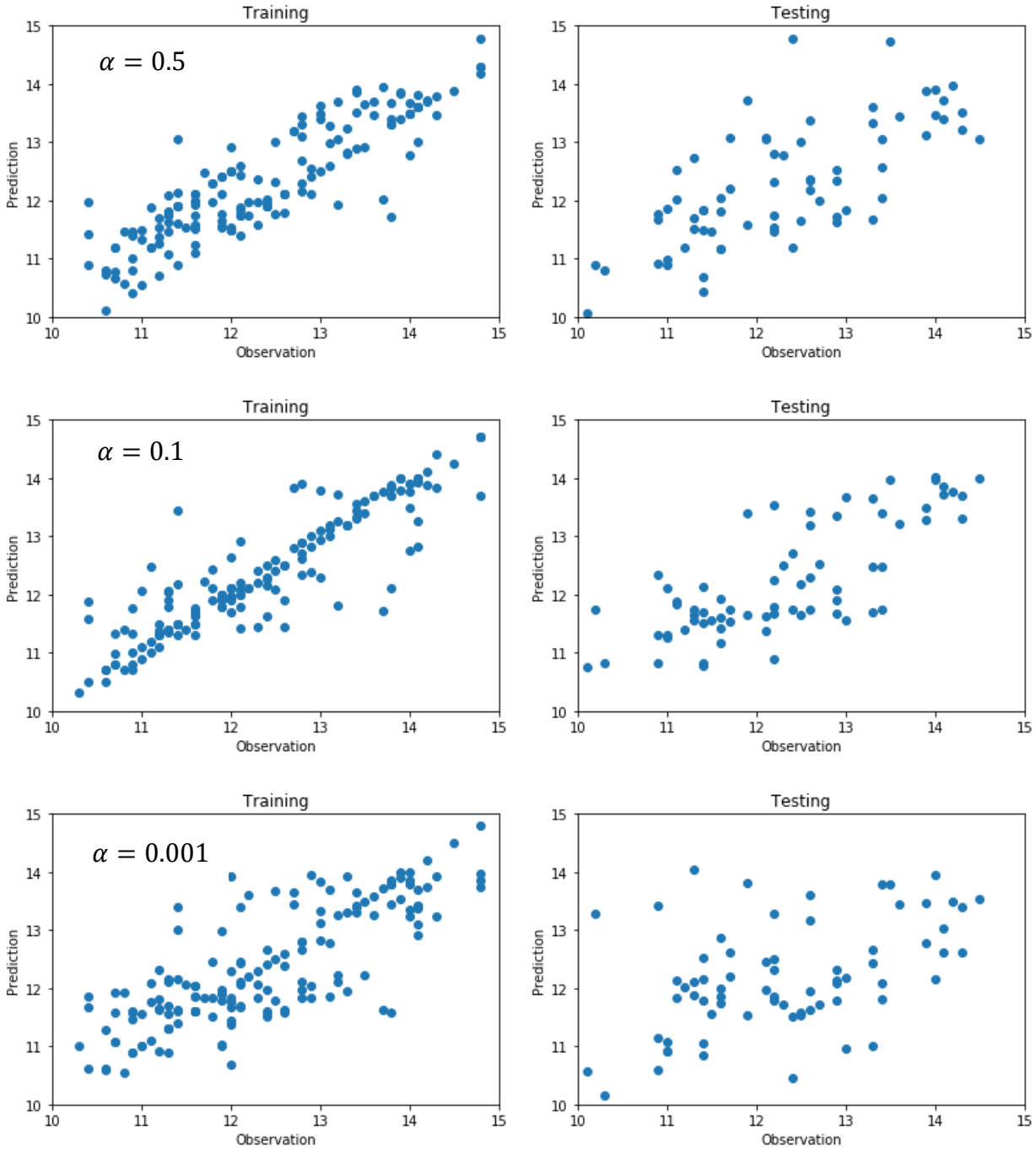

Figure 74 – Influence of the error factor on the model performances. Top graph: large error margin, Middle graph: "good" error margin. Lower graph: small error margin.

Appendix 4.1. Methodology

Note: The Mymensingh province is a new province since 2015, however, the administrative dataset used on GEE still includes the Mymensingh province in Dhaka, and we decided to keep this group of provinces together when displaying the results and validations.

Appendix 4.3. Model's performances

Figure 75 displays the difference in land cover between the Dynamic Land Cover (DLC) and the Random Forest Classification. The DLC was restricted to open water, buildings, crops, and other vegetation respectively in blue, grey, yellow, and green and is provided annually, meaning that crops encompass different types of crops and from different growing seasons. Water bodies are represented in blues and rivers as well as small reservoirs and ponds were correctly captured. Trees and buildings were also correctly estimated but small roads were missing. Rice areas were also correctly captured although riverbanks were sometimes misclassified as rice paddies due to their intermittent flooding patterns similar to flooded rice paddies (see Figure 76).

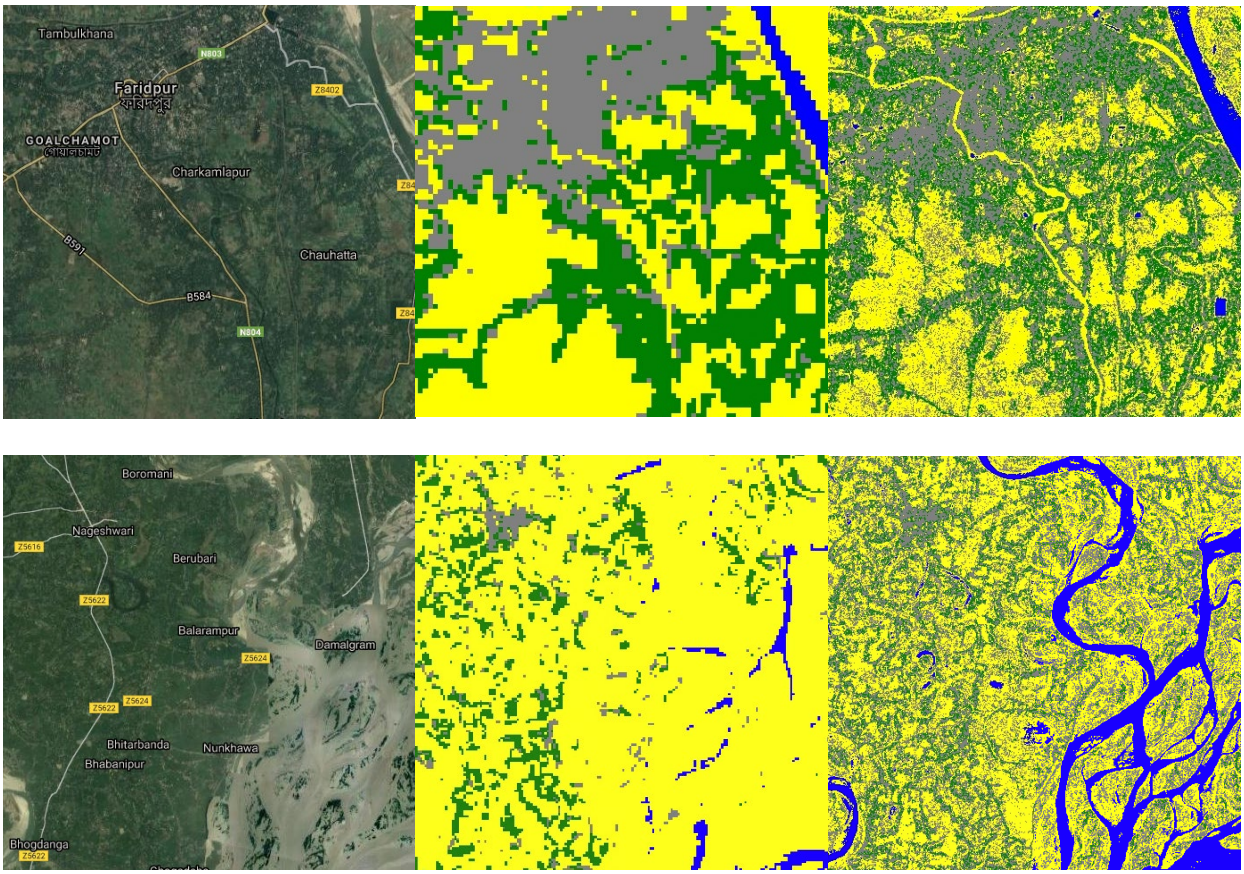

Figure 75 – Comparison of satellite imagery, Dynamic Land Cover, and our results at two levels of altitudes. Left: RGB Imagery from Google Earth over the Rangpur province. Middle: Dynamic Land Cover with water, buildings, crops, and other vegetations respectively blue, grey, yellow, and green. Right: Results of our classification with water, buildings, rice, and other vegetation respectively in blue, grey, yellow, and green.

Figure 76 illustrates one example of classification and misclassification of the land cover over Bangladesh. Riverbanks are classified as rice paddies even if it is unlikely that they are cultivated. Flooded areas can easily be taken out of rice paddies classifications by adding a vegetation index, such as NDVI, but grass growing on the riverbanks during periods of low flow and being flushed away when the river gains a higher flow are impossible to differentiate from rice paddies.

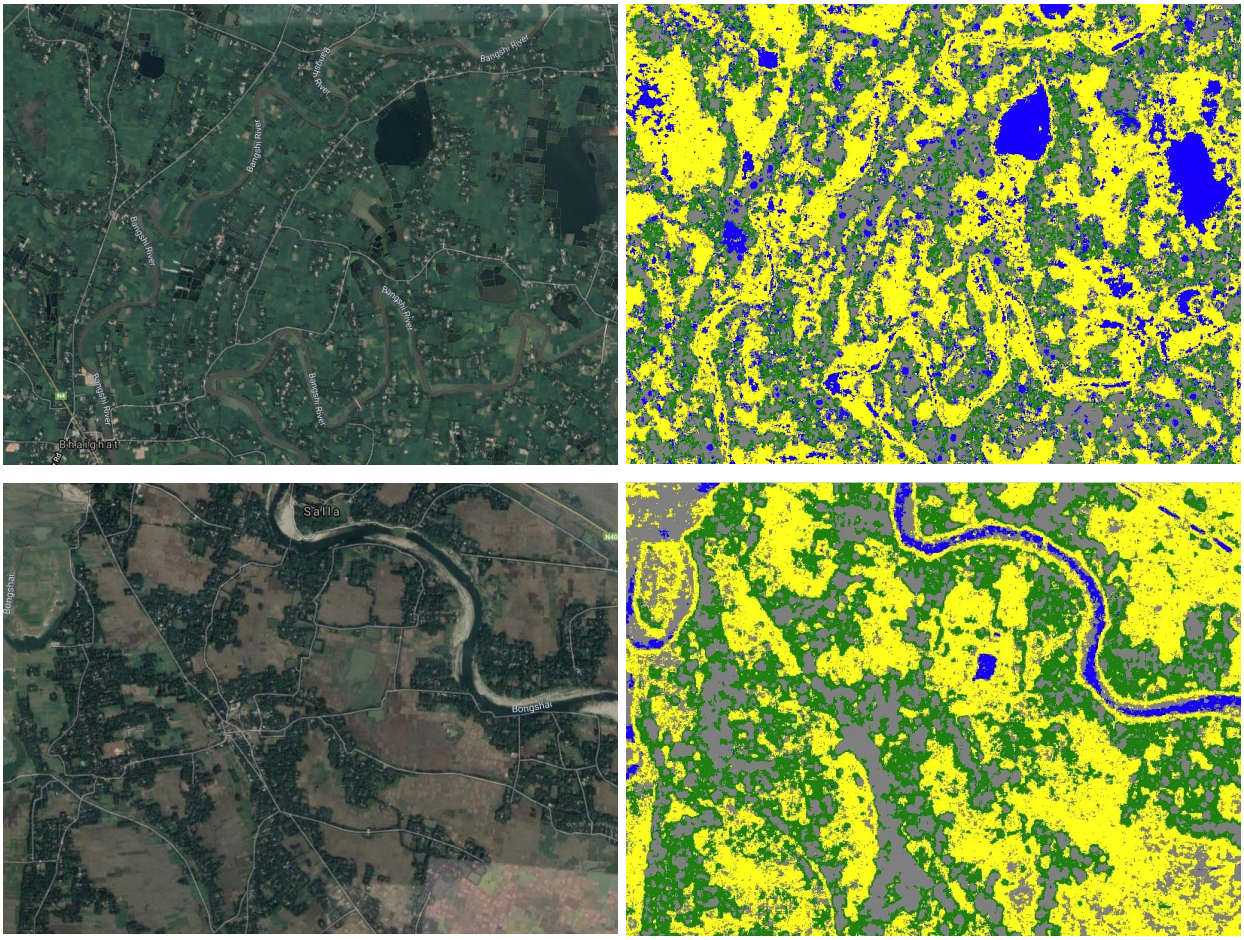

Figure 76 – Visual validation of the classification results for two different areas (Top and bottom plots). Left: RGB imageries in Bangladesh from Google Earth. Right: Results of the classification algorithm with rice paddies in yellow, tress in green, water bodies in blue, and buildings in grey.

Appendix 4.4. Additional results

Figure 77 indicates the yields of Aman rice by district (left map) and summarized by province (right chart).

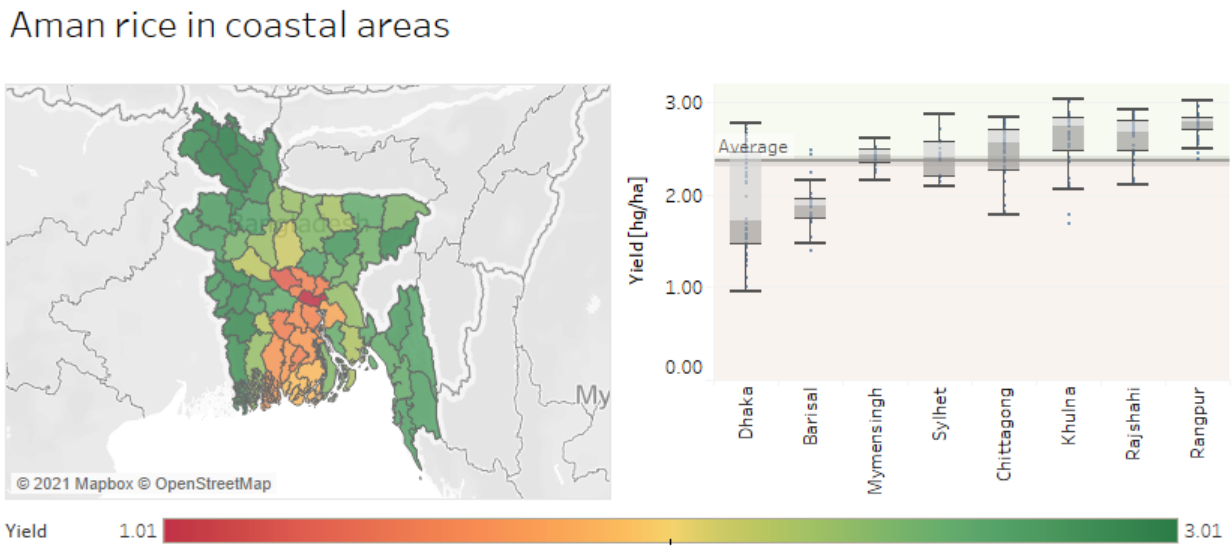

Figure 77 – Aman rice yields by districts in Bangladesh. The district in coastal areas exhibits a yield lower than the average, mainly due to high salinity and waterlogged croplands,

Changes in harvested area and yield at the district level obtained from governmental statistics are summarized in Figure 78.

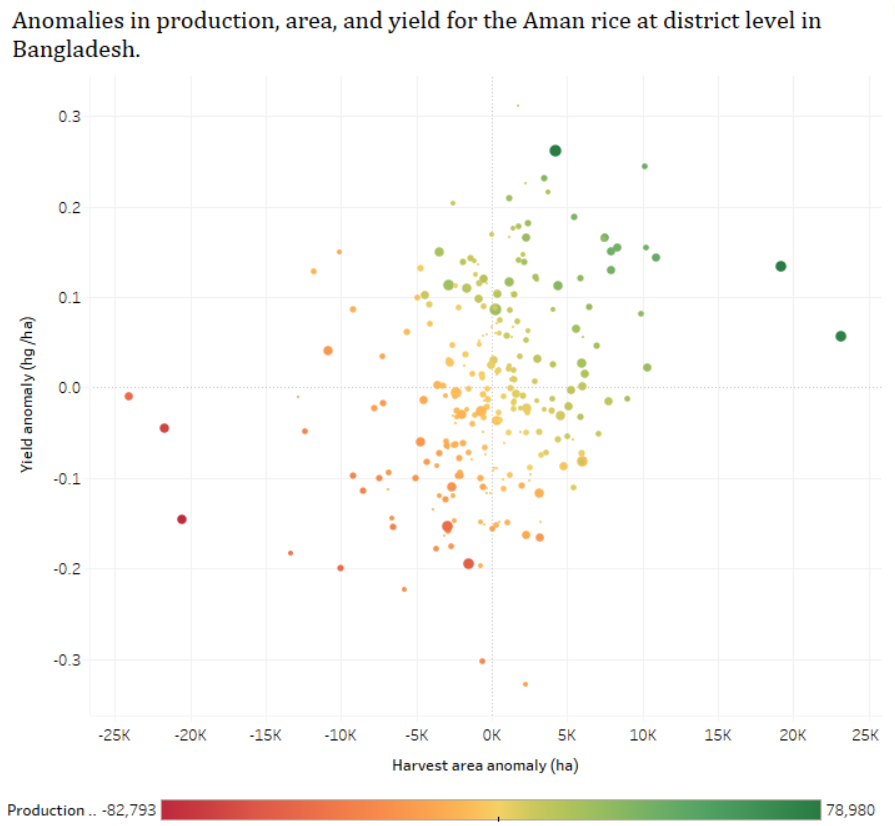

Figure 78 – Anomalies in yield, harvested area, and production at districts level in Bangladesh. The color indicates the change in production, the size of the dot indicates the total production.

REFERENCES

Adams, R., Rosenzweig, C., Peart, R., Ritchie, J., McCarl, B., Glyer, K., . . . Allen, L. (1990). Global climate change and United-States agriculture. *Nature*, 219-224.

AghaKouchak, A., Farahmand, A., Melton, F., Teixeira, J., Anderson, M., Wardlow, B., & C., H. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*.

Agrawal, M., Singh, B., Rajput, M., Marshall, F., & Bell, J. (2003). Effect of air pollution on peri-urban agriculture: a case study. *Environmental Pollution*, 323-329.

Agricultural and Agri-Food Canada. (2002). *Drought Watch.* Retrieved from http://www.agr.gc.ca/pfra/drought/maps/archives/e1020830.pdf

Ainsworth , E., & Long, S. (2004). What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. *New Phytologist*, 351-372.

Alauddin, M., & Sharma, B. (2013). Inter-district rice water productivity differences in Bangladesh: An empirical exploration and implications. *Ecological Economics*, 210-218.

Alexandratos, N., & Bruinsma, J. (2012). *World agriculture towards 2030/2050: the 2012 revision.* Rome: FAO.

Allen, F., & Gale, D. (2000). Financial contagion. *J. Polit. Econ.*, 1-33.

Avnery, S., Mauzerall, D., Liu, J., & Horowitz, L. (2001). Global crop yield reductions due to surface ozone exposure: 1. Year 2000 crop production losses and economic damage. *Atmospheric Environment*, 2284-2296.

Baggio, J., BurnSilver, S., Arenas, A., Magdanz, J., Kofinas, G., & De Domenico, M. (2016). Multiplex social ecological network analysis reveals how social changes affect community robustness more than resource depletion. *PNAS*, 13708-13713. doi:https://doi.org/10.1073/pnas.1604401113

Banerjee, L. (2010). Effects of Flood on Agricultural Productivity in Bangladesh. *Oxford Development Studies*.

Bangladesh Bureau of Statistics. (2021, 05 12). *Agriculture Wing.* Retrieved from Bangladesh Bureau of Statistics: http://bbs.portal.gov.bd

Battiston, S., Gatti, D., Gallegati, M., & Greenwald, B. (2012). Liaisons dangereuses: increasing connectivity, risk sharing, and systemic risk. *J. Econ. Dyn. Control*, 1121-1141.

Belton, P., & Taylor, J. (2004). Sorghum and millets: protein sources for Africa. *Trends in Food Science & Technology*, 94-98.

Benzinger, G., Kyle, J., & Blumenthal, K. (1998). A Specific Interaction between the Cardiac Sodium Channel and Site-3 Toxin Anthopleurin B. *The Journal of Biological Chemistry*, 80-84.

Berrittella, M., Hoekstra, A., Rehdanz, K., Roson, R., & Tol, R. (2007). The economic impact of restricted water supply: a computable general equilibrium analysis. *Water Research*, 1799–1813.

Bhattacharya, J., Saha, N., Mondal, M., Bhandari, H., & Humphreys, E. (2019). The feasibility of high yielding aus-aman-rabi cropping systems in the polders of the low salinity coastal zone of Bangladesh. *Field Crops Research*, 33-46.

Bollobas, B. (2001). *Random Graph.* Cambridge Studies in Advanced Mechanics.

Boser, B., Guyon, I., & Vapnik, V. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory* , 144.

Bren d'Amour, C., Wenz, L., Kalkuhl, M., Steckel, J., & Creutzig, F. (2016). Teleconnected food supply shocks. *Environmental Research Letters*.

Brown, M., Gunn, S., & Lewis, H. (1999). Support vector machines for optimal classification and spectral unmixing. *Ecological Modelling*, 167-179.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., . . . Fritz, S. (2020). Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe. *[Data set] Zenodo*. doi:http://doi.org/10.5281/zenodo.3939050

Burling, K., Hunsche, M., & Noga, G. (2011). Use of blue–green and chlorophyll fluorescence measurements for differentiation between nitrogen deficiency and pathogen infection in winter wheat. *Journal of Plant Physiology*, 1641-1648.

Cai, Y., Lin, H., & Zhang, M. (2019). Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Advances in Space Research*, 2233-2244.

Calzadilla, A., Rehdanz, K., & Tol, R. (2010). The economic impact of more sustainable water use in agriculture: a computable general equilibrium analysis. *J. Hydrol*, 292-305.

Casagrande, M., David, C., Valantin-Morison, M., Makowski, D., & Jeuffroy, M. (2009). Factors limiting the grain protein content of organic winter wheat in south-eastern France: a mixed-model approach. *Agronomy for Sustainable Development*, 565-574.

Casanova, D., Goudriaan, J., Catala Forner, M., & Withagen, J. (2002). Rice yield prediction from yield components and limiting factors. *European Journal of Agronomy*, 41-61.

Chameides, W., Yu, H., Liu, S., Bergin, M., Zhou, X., Mearns, L., . . . Giorg, F. (1999). Case study of the effects of atmospheric aerosols and regional haze on agriculture: An opportunity to enhance crop yields in China through emission controls? *Proceedings of the National Academy of Sciences*, 13626-13633.

Chiang, F., Massdiyasni, O., & AghaKouchak, A. (2021). Evidence of anthropogenic impacts on global drought frequency, duration, and intensity. *Nature Communications*.

Chiang, F., Mazdiyasni, O., & AghaKouchak, A. (2018). Amplified warming of droughts in southern United States in observations and model simulations. *Sciences Advances*.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogee, J., Allard, V., . . . Friend, A. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 528-533.

COPA-COGECA. (2003). *Assessment of the impact of the heat wave and drought of the summer 2003 on agriculture and forestry.*

CRED. (2019). *The International Disaster Database - EM-DAT.* Louvain: Centre for Research on the Epidemiology of Disasters.

Cruz, L., Cagampang, G., & Juliano, B. (1970). Biochemical Factors Affecting Protein Accumulation in the Rice Grain. *Plant Physiology*.

Dalla Marta, A., Grifoni, D., Mancini, M., Zipoli, G., & Orlandini, S. (2011). The influence of climate on durum wheat quality in Tuscany, Central Italy. *International Journal of Biometeorology*, 87-96.

Daryanto, S., Wang, L., & Jacinthe, P.-A. (2015). Global Synthesis of Drought Effects on Food Legume Production. *PLOS ONE*. doi:https://doi.org/10.1371/journal.pone.0127401

Dasgupta, S., Hossain, M., Huq, M., & Wheeler, D. (2014). Climate change and soil salinity: The case of coastal Bangladesh. *Ambio*, 815–826.

Dasgupta, S.; Hossain, M.; Huq, M.; Wheeler, D. (2014). *Facing the hungry tide : climate change, livelihood threats, and household responses in coastal Bangladesh (English).* Washington, D.C: World Bank Group.

De Bono, A., Peduzzi, P., Kluser, S., & Giuliani, G. (2004). Impacts of Summer 2003 Heat Wave in Europe. *United Nations Environment Programme, Environment Alert Bulletin*.

Ding, Y., Hayes, M., & Widhalm, M. (2011). Measuring economic impacts of drought: a review and discussion. *Disaster Prevention and Management*, 434-446.

Distefano, T., Laio, F., Ridolfi, L., & Shiavo, S. (2018). Shock transmission in the International Food Trade Network. *PLOS ONE*.

Drewnowski , A., & Popkin, B. (1997). The nutrition transition: New trends in the global diet. *Nutrition review*, 31-43.

Dror, Y., Rimon, E., & Vaida, R. (2020). The Kernel Organs and Composition. *Whole-Wheat Bread for Human Health*, 65-90.

Dupont, F., & Altenbach, S. (2003). Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *Journal of cereal science*, 133-146.

Ercsey-Ravasz, M., Toroczkai, Z., Lakner, Z., & Baranyi, J. (2012). Complexity of the international agro-food trade network and its impact on food safety. *PLoS ONE*.

Fagiolo, G., Reyes, J., & Schiavo, S. (2010). The evolution of the world trade web: a weighted-network analysis. *J. Evol. Econ.*, 479-514.

Fang, H. (1988). Rice crop area estimation of an administrative division in China using remote sensing data. *International Journal of Remote Sensing*, 3411-3419.

Fangmeier, A., Temmerman, L., Mortensen, L., Kemp, K., Burke, J., Mitchell, R., . . . Weigel, H. (1999). Effects on nutrients and on grain quality in spring wheat crops grown under elevated CO2 concentrations and stress conditions in the European, multiple-site experiment 'ESPACE-wheat. *European Journal of Agronomy*, 215-229.

FAO. (2011). *Food Export Restrictions: Review of the 2007-2010 Experience and Considerations for Disciplining Restrictive.* FAO Commodity and Trade Policy Research.

FAO. (2012). *Coping with water scarcity - An action framework for agriculture and food security.* Rome: Food and Agriculture Organization of the United Nations.

FAO. (2020). Annex 1 - FAO work on statistics. In FAO, *World Food and Agriculture - Statistical Yearbook 2020.* Rome, Italy: FAO.

FAO/WHO/UNU. (2001). *Human energy requirements. Report of a joint FAO/WHO/UNU expert consultation.* Rome, Italy: Food and Agriculture Organization of the United Nations.

Feng, Z., & Kobayashi, K. (2009). Assessing the impacts of current and future concentrations of surface ozone on crop yield with meta-analysis. *Atmospheric Environment*, 1510–1519.

Food and Agriculture Organization. (2021). *Crop Calendar dataset.* Retrieved from http://www.fao.org/agriculture/seed/cropcalendar/welcome.do

Freeman, K., Raun, W., Johnson, G., Mullen, R., Stone, M., & Solie, J. (2003). Late-season prediction of wheat grain yield and grain protein. *Communications in Soil Science and Plant Analysis*, 1837-1852.

Fuhrer, J., Skarby, L., & Ashmore, M. (1997). Critical levels for ozone effects on vegetation in Europe. *Environmental pollution*, 91-106.

Fukuda, S., & Hirosawa, H. (2001). Support vector machine classification of land cover: application to polarimetric SAR data. *International Geoscience and Remote Sensing Symposium* (pp. 187-189). IEEE.

Gai, P., & Kapadia, S. (2010). Contagion in financial networks. *Proc. R. Soc. A*, 2401-2423.

Gaju, O., Allard , V., Martre, P., Snape, J., Le GOuis, J., MOreau, D., . . . Foulkes, M. (2011). dentification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crops Research*, 139-152.

Gamuyao, R., Chin, J., Pariasca-Tanaka, J., Pesaresi, P., Catausan, S., Dalid, C., . . . Heuer, S. (2012). The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature*, 535-539. doi:10.1038/nature11346

Garlaschelli, D., & Loffredo, M. (2005). Structure and evolution of the world trade network. *Physica A: Statistical Mechanics and its Applications*, 138-144.

Ghosh, J. (2009). The Unnatural Coupling: Food and Global Finance. *Journal of Agrarian Change*.

Godfray, C., Beddington, J., Crute, I., Haddad, L., Lawrence, D., Muir, J., . . . Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*.

Grantz, D., Garner, J. H., & Johnson, D. (2003). Ecological effects of particulate matter. *Environment International*, 213-239.

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J., . . . Griffis, T. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *PNAS*, 1327-1333.

Hamaker, B. (1994). The influence of rice protein on rice quality. In W. Marshall, & J. Wadsworth, *Rice Science and Technology* (pp. 177-194). New York.

Hansen, P., Jorgensen, J., & Thomsen, A. (2002). Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression. *Journal of Agricultural Science*, 307-318.

Hao, Z., & AghaKouchak, A. (2014). A Nonparametric Multivariate Multi-Index Drought Monitoring Framework. *Advancing Drought Monitoring and Prediction*, 89-101.

Haque, S. (2006). Salinity Problems and crop production in coastal regions of Bangladesh. *Pakistan Journal of Botany*, 1359-1365.

Hatfield, J., & Prueger, J. (2015). Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes*, 4-10.

Hobbs, P., & Mehla, R. (2003). *The problem of late planting in wheat. Addressing resource conservation issues in rice–wheat systems of south Asia: a resource book.* (RWC-CIMMYT, Ed.) New Delhi, India: Rice–Wheat Consortium for the Indo-Gangetic Plains – International Maize and Wheat Improvement Center.

Hogy, P., & Fangmeier, A. (2008). Effects of elevated atmospheric CO2 on grain quality of wheat. *Journal of Cereal Science*, 580-591.

Hopkins, C. (1899). Improvement in the Chemical Composition of the Corn Kernel. *Journal of the American Chemical Society*, 1039–1057.

Hossain, M. (2006). Arsenic contamination in Bangladesh—An overview. *Agriculture, Ecosystems & Environment*, 1-16.

Howe, P., Mildenberger, M., Marlon, J., & Leiserowitz, A. (2015). Geographic variation in opinions on climate change at state and local scales in the USA. *Nature Climate Change*, 596–603.

Huang, J., Wang, X., Li, X., Tian, H., & Pan, Z. (2013). Remotely Sensed Rice Yield Prediction Using Multi-Temporal NDVI Data Derived from NOAA's-AVHRR. *PLoS ONE*.

Huang, Z., Vodenska, I., Havlin, S., & Stanley, H. (2013). Cascading Failures in Bi-partite Graphs: Model for Systemic Risk Propagation. *Nature Scientific Reports*.

IMPACT Development Team. (2012). *International Model for Polic Analysis of Agricultural Commodities and Trade (IMPACT): Model Description.* Washington, DC: International Food Policy Research Institute.

International Labour Organization. (2019). *"Employment by sex and age — ILO modelled estimates.".*

IRRI. (2021). *International Rice Research Institute.*

Jarvis, C., Sapirstein, H., Bullock, P., Naeem, H., Angadi, S., & Hussain, A. (2008). Models of growing season weather impacts on breadmaking quality of spring wheat from producer fields in western Canada. *Journal of the Science of Food and Agriculture*, 2357–2370.

Jia, M., Rossini, M., Colombo, R., & Celesti, M. (2021). Estimation of leaf nitrogen content and photosynthetic nitrogen use efficiency in wheat using sun-induced chlorophyll fluorescence at the leaf and canopy scales. *European Journal of Agronomy*, 122-126.

Joachims, T. (2002). *Learning to classify text using support vector machines: methods, theory and algorithms.* Springer.

Johansson, E., & Svensson, G. (1998). Variation in bread-making quality: Effects of weather parameters on protein concentration and quality in some Swedish wheat cultivars

grown during the period 1975–1996. *Journal of Science of Food and Agriculture*, 109-118.

Jones, D., Cross, P., Withers, P., DeLuca, T., Robinson, D., Quilliam, R., . . . Edwards-Jones, G. (2013). Nutrient stripping: the global disparity between food security and soil nutrient stocks. *Journal of Applied Ecology*. doi:10.1111/1365-2664.12089

Kanti Paul, B., & Rasid, H. (1993). Flood Damage to Rice Crop in Bangladesh. *Geographical Review*, 150-159.

Karlsson, P., Braun, S., Broadmeadow, M., Elvira, S., Emberson, L., Gimeno, B., . . . Wilkinson, M. (2007). Risk assessments for forest trees: The performance of the ozone flux versus the AOT concepts. *Environmental pollution*, 608-616.

Khoury, C., Bjorkman, A., Dempewolf, H., Ramirez-Villegas, J., Guarino, L., Jarvis, A., . . . Struik, P. (2014). Increasing homogeneity in global food supplies and. *PNAS*, 4001-4006.

Kichey, T., Hirel, B., Heumez, E., Dubois, F., & Le Gouis, J. (2007). In winter wheat (Triticum aestivum L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. *Field Crops Research*, 22-32.

Kimball, S., & Idso, S. (1983). Increasing Atmospheric CO2: Effects on Crop Yield, Water use and Climate. *Developments in Agricultural and Managed Forest Ecology*.

Konar, M., Lin, X., Ruddell, B., & Sivapalan, M. (2018). Scaling properties of food flow networks. *PLOS ONE*. doi:https://doi.org/10.1371/journal.pone.0199498

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., . . . Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications. *Atmospheric Chemistry and Physics*, 417-432.

Laborte, A., Gutierrez, M., Balanza, J., Saito, K., Zwart, S., Boschetti, M., . . . Nelson, A. (2017). RiceAtlas, a spatial database of global rice calendars and production. *Nature Scientific Data*.

Lang, T., & Barling, D. (2012). Food security and food sustainability: reformulating the debate†. *The Geographical Journal*.

Lawrence, G., Lyons, K., & Wallington, T. (2010). Introduction: Food Security, Nutrition and Sustainability in a Globalized World. In G. Lawrence, K. Lyons, & T. Wallington, *Food security, Nutrition, and Sustainability* (p. 23).

Le Toan, T., Ribbes, F., Wang, L., Floury, N., Ding, K., Kong, J., . . . Kurosu, T. (1997). Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. *IEEE Transaction on geoscience and remote sensing*, 41-56.

Li, X., & Xiao, J. (2019). A global, 0.05-degree product of solar-induced chlorophyll fluorescence. *Remote Sensing*, 517.

Li, X., Wu, L., Geng, X., Xia, X., Wang, X., Xu, Z., & Zu, Q. (2018). Deciphering the Environmental Impacts on Rice Quality for Different Rice Cultivated Areas. *Rice*.

Li-Hong, X., Wei-Xing, C., & Lin-Zhang, L. (2007). Predicting Grain Yield and Protein Content in Winter Wheat at Different N Supply Levels Using Canopy Reflectance Spectra. *Pedosphere*, 646-653.

Lin, X., Ruess, P., Marston, L., & Konar, M. (2019). Food flows between counties in the United States. *Environmental Research Letters*.

Liu, L., Wang, J., Bao, Y., Huang, W., Ma, Z., & Zhao, C. (2006). Predicting winter wheat condition, grain yield and protein content using multi-temporal EnviSat-ASAR and Landsat TM satellite images. *International Journal of Remote Sensing*, 737-753.

Lobell, D. (2007). Changes in diurnal temperature range and national cereal yields. *Agricultural and Forest Meteorology*, 229-238.

Lobell, D., & Gourdji, S. (2012). The influence of climate change on global crop productivity. *Plant Physiology*.

Lowder, S., Skoet, J., & Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development*, 16-29.

Magney, T., Eitel, J., Huggin, D., & Vierling, L. (2016). Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality. *Agricultural and Forest Meteorology*, 44-60.

Magney, T., Eitel, J., Huggins, D., & Vierling, L. (2016). Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality. *Agricultural and Forest Meteorology*, 46-60.

Mansaray, L., Wang, F., Huang, J., Yang, L., & Kanu, A. (2020). Accuracies of support vector machine and random forest in rice mapping with Sentinel-1A, Landsat-8 and Sentinel-2A datasets. *Geocarto International*, 1088-1108.

Mathur, A., & Foody, G. (2004). Land cover classification by support vector machine: towards efficient training. *International Geoscience and Remote Sensing Symposium* (pp. 742-744). IEEE.

Matre, P., Jamieson, P., Semenov, M., Zyskowski, R., Porter, J., & Triboi, E. (2006). Modelling protein content and composition in relation to crop nitrogen dynamics for wheat. *European Journal of Agronomy*, 138-154.

Mazdiyasni, O., & AghaKouchak, A. (2015). Substantial increase in concurrent droughts and heatwaves in the United States. *Proceedings of the National Academy of Sciences of the United States*, 11484 - 11489.

McCool, P., Musselman, R., & Teso, R. (1986). *Determining crop yield losses from air pollutants.* California Agriculture.

McDaniel, C., & Balistreri, E. (2003). A review of Armington trade substitution elasticities. *Economie Internationale*.

McKee, T., Doesken, N., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference of Applied Climatology*, 179–184.

Medek, D., Schwartz, J., & MyersqS. (2017). stimated Effects of Future Atmospheric CO2 Concentrations on Protein Intake and the Risk of Protein Deficiency by Country and Region. *Environmental Health Perspectives*.

Medellín-Azuara, J., MacEwan, D., Howitt, R., Sumner, D., & Lund, J. (2016). *Economic Analysis of the 2016 California Drought on Agriculture. A report for the California Department of Food and Agriculture.* Davis, California: Center for Watershed Sciences University of California Davis.

Meyfroidt, P. (2017). Mapping farm size globally: benchmarking the smallholders debate. *Environmental Research Letters*.

Monks, P., Granier, C., Fuzzi, S., Stohl, A., Williams, M., Akimoto, H., . . . Blake, R. (2009). Atmospheric composition change – global and regional air quality. *Atmospheric Environment*, 5268-5350.

Moschini, G., & Hennessy, A. (2001). Uncertainty, risk aversion, and risk management for agricultural producers. In B. Gardner , & G. Rausser, *Agricultural Production* (pp. 87-153).

Myers, S., Wessells, K., Kloog, I., Zanobetti, A., & Schwartz, J. (2015). Effect of increased concentrations of atmospheric carbon dioxide on the global threat of zinc deficiency: a modelling study. *Lancet Global Health*.

Myers, S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A., Carlisle, B., . . . Dietterich, L. (2014). Increasing CO2 threatens human nutrition. *Nature*, 139–142.

NASA Earth Observatory. (2009). *Drought in Argentina.* NASA. Retrieved from https://earthobservatory.nasa.gov/images/37105/drought-in-argentina

Nasim, M., Shahidullah, S., Saha, A., Muttaleb, M., Aditiya, T., Ali, M., & Kabir, M. (2017). Distribution of Crops and Cropping Patterns in Bangladesh. *Bangladesh Rice Journal*, 1-55.

Nelson, G., Valin, H., Sands, G., Havlík, P., Ahammad, H., Deryng, D., . . . Hasegawa, T. (2013). Climate change effects on agriculture: Economic responses to biophysical shocks. *PNAS*.

Neumann, E., Ashfaque, K., Badruzzaman, A., Ashraf Ali, M., Shoemaker, J., & Harvey, C. (2009). Anthropogenic influences on groundwater arsenic concentrations in Bangladesh. *Nature Geoscience*, 46-52. doi:doi.org/10.1038/ngeo685

Oatley, T., Winecoff, K., Pennock, A., & Danzman, S. (2013). The Political Economy of Global Finance: A Network Model. *Perspectives on Politics*, 133-153.

Okamoto, K., & Fukuhara, M. (1996). Estimation of paddy rice field area using the area ratio of categories in each pixel of Landsat TM. *International Journal of Remote Sensing*, 1735-1749.

Otero, G., Pechlaner, G., & Gurcan, E. (2013). The Political Economy of "Food Security" and Trade: Uneven and Combined Dependency. *Rural Sociology*.

Ozturk, A., & Aydin, F. (2004). Effect of Water Stress at Various Growth Stages on Some Quality Characteristics of Winter Wheat. *Journal of Agronomy and Crop Science*, 93-99.

Pan, J., Zhu, Y., Cao, W., Dai, T., & Jiang, D. (2006). Predicting the protein content of grain in winter wheat with meteorological and genotypic factors. *Plant Production Science*, 323-333.

Panozzo, J., & Eagles, H. (2000). Cultivar and environmental effects on quality characters in wheat. II. Protein. *Australian Journal of Agriculture Research*.

Panozzo, J., Walker, C., Partington, D., Neumann, N., Tausz, M., Seneweera, S., & Fitzgerald, G. (2014). Elevated carbon dioxide changes grain protein concentration and composition and compromises baking quality. A FACE study. *Journal of Cereal Science*, 461-470.

Park, S., Im, J., Park, S., Yoo, C., Han, H., & Rhee, J. (2018). Classification and Mapping of Paddy Rice by Combining Landsat and SAR Time Series Data. *Remote Sensing*.

Parry, M., Rosenzweig, C., & Livermore, M. (2005). Climate change, global food supply and risk of hunger. *Philosophical Transactions of the Royal Society*, 2125-2138.

Peng, D., Huete, A., Huang, J., Wang, F., & Sun, H. (2011). Detection and estimation of mixed paddy rice cropping patterns with MODIS data. *International Journal of Applied Earth Observation and Geoinformation*, 13-23.

Pirgozliev, V., Birch, C., Rose, S., Kettlewell, S., Bedford, P., & Bedford M. (2003). Chemical composition and the nutritive quality of different wheat cultivars for broiler chickens. *British Poultry Science*, 464-475.

Pleijel, H., Karlsson, G., Silb, E., Danielsson, H., Skarby, L., & Sellden, G. (1996). Exposure of a grass-clover mixture to ozone in open-top chambers—effects on yield, quality and botanical composition. *Agriculture, Ecosystems & Environment*, 55-62.

Porter, J., & Semenov, M. (2005). Crop responses to climatic variation. *Philosophical Transactions of the Royal Society*, 2021–2035.

Portmann, F., Siebert, S., & Döll, P. (2010). MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modelling. *Global Biochemical Cycles*.

Prosekov, A., & Ivanova, S. (2018). Food security: The challenge of the present. *Geoforum*, 73-77.

Puri, S., Dhillon, B., & Sodhi, N. (2014). Effect of Degree of Milling (Dom) on Overall Quality of Rice - A Review. *International Journal of Advanced Biotechnology and Research*, 474-489.

Rahman, A., Kang, S., Naghabhatla, N., & Macnee, R. (2017). Impacts of temperature and rainfall variation on rice productivity in major ecosystems of Bangladesh. *Agriculture and Food Security*. doi:10.1186/s40066-017-0089-5

Randall, P., & Moss, H. (1990). Some effects of temperature regime during grain filling on wheat quality. *Australian journal of agricultural research*.

Rao, A., Smith, J., Jandhyala, V., Papendick, R., & Par, J. (1993). Cultivar and Climatic Effects on the Protein Content of Soft White Winter Wheat. *Agronomy Journal*, 1023-1028.

Reilly, M., & Willenbockel, D. (2010). Managing uncertainty: a review of food system scenario analysis and modelling. *Food Policy*.

Reilly, T., & Kane, S. (1992).) Economic implications of global climate change for world agriculture. *J Agric Resour Econ*, 195-204.

Rhoades, S. (1993, March). The Herfindahl-Hirschman Index. *Federal Reserve Bulletin*.

Rmalli, S., Jenkins, R., Watts, M., & Haris, P. (2012). Reducing human exposure to arsenic, and simultaneously increasing selenium and zinc intake, by substituting non-aromatic rice with aromatic rice in the diet. *Biomedical Spectroscopy and Imaging*. doi:10.3233/BSI-120028

Rosegrant, M., Ringler, C., Msangi, S., Sulser, T., Zhu, T., & Cline, S. (2008). *International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): Model Description.* Washington, D.C.: International Food Policy Research Institute.

Rosenzweig, C., & Parry, M. (1994). Potential impact of climate change on world food. *Nature*, 133-138.

Rulli, M., & D'odorico, P. (2013). The fourth food revolution. *Nature Geoscience*, 417-418.

Sacks, W., Deryng, D., Foley, J., & Ramankutty, N. (2010). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*.

Serrano, M., & Boguna, M. (2003). Topology of the world trade web. *Phys. Rev. E*.

Sgroi, L., Lovino, M., Berbery, E., & Müller, G. (2021). Characteristics of droughts in Argentina's core crop region. *Hydrology and Earth System Sciences*, 2475–2490.

Sheema, M., Ali, L., & Akhtar, M. (2002). Wheat following the cotton and rice-based cropping system in Pakistan. *Annual Wheat Newsletter*, pp. 112-113.

Smika, D., & Greb, B. (1973). Protein Content of Winter Wheat Grain as Related to Soil and Climatic Factors in the Semiarid Central Great Plains1. *Agronomy Journal*.

Smika, D., & Greb, B. (1973). Protein Content of Winter Wheat Grain as Related to Soil and Climatic Factors in the Semiarid Central Great Plains1. *Agronomy Journal*, 433-436.

Smith, A., & Katz, R. (2013). US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. *Natural Hazards*, 387-410.

Song, L., Guanter, L., Guan, K., You, L., Huete, A., Ju, W., & Zhang, Y. (2018). Satellite sun-induced chlorophyll fluorescence detects early response of winter wheat to heat stress in the Indian Indo-Gangetic Plains. *Global Change Biology*, 4023-4037.

Spahni, R., Wania, R., Neef, L., Van Weele, M., Pison, I., Bousquet, P., . . . Van Velthoven, P. (2011). Constraining global methane emissions and uptake by ecosystems. *Biogeosciences*, 1643-1665.

Steenfeldt, S. (2001). The dietary effect of different wheat cultivars for broiler chickens. *British Poultry Science*, 595-609.

Sutton, M., Amann, M., Emberson, L., & Dentener, F. (2015). *Sustainable Agriculture and Air Pollution : reducing emissions leads to many benefits. Event at the Milan World Expo.* Milano, Italia.

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., . . . Miskus, D. (2002, August). The drought monitor. *Bulletin of the American Meteorological Society.*, pp. 1181-90.

Tai, A., Martin, M., & Heald, C. (2014, July 24). Threat to future global food security from climate change and ozone air pollution. *Nature*.

Tan, C.-P., Koay, J., Lim, K., Ewe, H., & Chuah, H. (2007). Classification of Multi-Temporal SAR Images for Rice Crops Using Combined Entropy Decomposition and Support Vector Machine Technique. *Progress In Electromagnetics Research*, 19-39.

Taylor, A., & Gilmour, A. (1971). Wheat protein prediction from climatic factors in southern NSW. *Australian Journal of Experimental Agriculture*, 546-549.

Teixeira, E., Fischer, G., van Velthuizen, H., van Dingenen, R., Dentener, F., Mills, G., . . . Ewert, F. (2011). Limited potential of crop management for mitigating surface ozone impacts on global food supply. *Atmospheric Environment*, 2569-2576.

Terman, G. (1979). Yields and Protein Content of Wheat Grain as Affected by Cultivar, N, and Environmental Growth Factors1. *Agronomy Journal*.

Tong, D., Mathur, R., Schere, K., Kang, D., & Yu, S. (2007). The use of air quality forecasts to assess impacts of air pollution on crops: Methodology and case study. *Atmospheric Environment*, 8772-8784.

Torbick, N., Chowdhury, D., Salas, W., & Qi, J. (2017). Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2. *Remote Sensing*.

Triboi, E., Martre, P., & Triboi-Blondel, A. (2003). Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *Journal of Experimental Botany*, 1731-1742.

Turner, M., & Congalton, R. (1998). Classification of multi-temporal SPOT-XS satellite data for mapping rice fields on a West African floodplain. *International Journal of Remote Sensing*, 21-41.

U.S. Department of Agriculture. (1957). *Distribution of the Varieties and Classes of Wheat in the United States in 1954.* Washington DC: U.S. Department of Agriculture.

UNEP. (2004, March). Impacts of summer 2003 heat wave in Europe. *Environment Alert Bulletin*.

US Grains Council. (2021). *2020/2021 Corn Harvest Quality Report.* US Grain Council.

USDA. (2016). *Brazil's Corn Industry and the Effect on the Seasonal Pattern of U.S. Corn Exports.* USDA Economic Research Service.

USDA Foreign Agricultural Service Office of Global Analysis. (2009). *World Agricultural Production.* United States Department of Agriculture.

van der Velde, M., Wriedt, G., & Bouraoui, F. (2010). Estimating irrigation use and effects on maize yield during the 2003 heatwave in France. *Agriculture, Ecosystems & Environment*, 90-97.

Van Dingenen, R., Dentener, F., Raes, F., Krol, M., Emberson, L., & Cofala, J. (2009). The global impact of ozone on agricultural crop yields under current and future air quality legislation. *Atmospheric Environment*, 604-618.

Van Niel, T., McVicar, T., Fang, H., & Liang, S. (2003). Calculating environmental moisture for per-field discrimination of rice crops. *International Journal of Remote Sensing*, 885-890.

Vlachokostas, C., Nastis, S., Achillas, C., Kalogeropoulos, K., Karmiris, I., Moussiopoulos, N., . . . Limperi, N. (2010). Economic damages of ozone air pollution to crops using combined air quality and GIS modelling. *Atmospheric Environment*, 3352-3361.

Vogel, E., & Meyer, R. (2018). Chapter 3 - Climate Change, Climate Extremes, and Global Food Production—Adaptation in the Agricultural Sector. *Resilience*, 31-49.

Wang, L., Tian, Y., Yao., X., Zhu, Y., & Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *Field Crops Research*, 178-188.

Wang, Z., Wang, J., Zhao, C., Zhao, M., Huang, W., & Wang, C. (2007). Vertical Distribution of Nitrogen in Different Layers of Leaf and Stem and Their Relationship with Grain Quality of Winter Wheat. *Journal of Plant Nutrition*, 73-91.

Wardlaw, I., Blumenthal, C., Larroque, O., & Wrigley, C. (2002). Contrasting effects of chronic heat stress and heat shock on kernel weight and flour quality in wheat. *Funct. Plant Biol*, 25-34.

Wargent, J., & Jordan, B. (2013). From ozone depletion to agriculture: understanding the role of UV radiation in sustainable crop production. *NewPhytologist*, 1058-1076.

Webb, P. (2009). Fiat Panis: Für eine Welt ohne Hunger. In H. Eiselen. Stuttgart: Hampp Media.

Wilkinson, J. (2009). The globalization of agribusiness and developing world food systems. *Monthly Reviews*.

World Bank. (2018). *World Bank Survey.* Washington DC.

World Health Organization. (2016). *Micronutrient Deficiencies.*

Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., . . . Moore, B. (2005). Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, 480-492.

Xie, L., Xhang, H., Wu, F., Wang, C., & Zhang, B. (2015). Capability of Rice Mapping Using Hybrid Polarimetric SAR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3812-3822. doi:10.1109/JSTARS.2014.2387214

Yan, X., Akiyama, H., Yagi, K., & Akmoto, H. (2009). Global estimations of the inventory and mitigation potential of methane emissions from rice cultivation conducted using the 2006 Intergovernmental Panel on Climate Change Guidelines. *Global Biogeochemical Cycles*.

Yao, H., & Conrad, R. (1999). Thermodynamics of methane production in different rice paddy soils from China, the Philippines and Italy. *Soil Biology and Biochemistry*, 463-473.

Yoshida, S. (1981). *Fundamentals of Rice Crop Science.* Manila, Philippines: The International Rice Research Institute.

Zhang, Y., Wang, C., Wu, J., Qi, J., & Salas, W. (2009). Mapping paddy rice with multitemporal ALOS/PALSAR imagery in southeast China. *International Journal of Remote Sensing*, 6301-6315.

Zhu, C., Kobayashi, K., Loladze, I., Zhu, J., Jiang, Q., Xu, X., . . . Ziska, L. (2018). Carbon dioxide (CO2) levels this century will alter the protein, micronutrients, and vitamin content of rice grains with potential health consequences for the poorest rice-dependent countries. *Science Advances, 4*(5).