USING CROP SIMULATION TO OPTIMIZE VARIABLE RATE
EXPERIMENTATION

BY

GERMAN MANDRINI

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Master's Committee:
      Professor David S. Bullock, Adviser
      Professor Taro Mieno, University of Nebraska - Lincoln
      Professor Nicolás A. Martín
      Professor Nicholas D. Paulson

ABSTRACT

Researchers working on a USDA-sponsored research project are exploring a new concept of on-farm experimentation (OFE). These trials are implemented by farmers at their fields, in a similar way to how they would plant a regular production crop. This concept generates large amounts of data at low cost that, after processing, will generate local models about the yield response function within a field.

At the time of this work, the research group is running more than 100 trials in different states and countries. There are questions related to how to optimize OFE. To address those questions, the APSIM crop growth model was used to simulate the concept of running on-field trials, use that information to calculate the Economic Optimum Nitrogen Rate (EONR), and finally use that EONR in a regular crop production. Spatially variable layers of data that characterized a field were transformed into APSIM parameters. Daily weather events were obtained from historical weather data for the field's county. Economic analysis of different strategies was performed, which involved testing if the increase in revenues due to including more variables or running more trials outperforms the cost, and how weather affects the results.

The results will help to optimize the actual protocol that is guiding the implementation of the trials. Key results obtained by this research were: (1) The value of conducting trials and using that information for N-management advice was 9.8 $/ha. (2) The added value of gathering soil sampling data at the same time was 7.4 $/ha. (3) The optimal time to stop running trials and start using the information for N-management advice was one or two years, depending on the weather. (4) Conducting trials and using that information for N-management advice decreased N-leaching by 10.4 kg/ha. Performing soil sampling tests together with running trials made N-management advice increase the efficiency and reduced N-leaching by 5.9 kg/ha more. (5) A tentative rule for deciding if a one trial year is sufficient or if one more year is needed was obtained by determining the likelihood of the weather of the trial year compared with the historic weather. These results provide insights that will be helpful to optimize the protocol that guide OFE and help farmers increase profits in the fastest way and decrease the environmental impact of nitrogen fertilization.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 EVOLUTION OF ON-FARM, LARGE SCALE FIELD TRIALS

In 1905, the Haber-Bosch process was invented, a process that industrially transforms nitrogen gas, abundant in the atmosphere, into ammonia, which could be absorbed and used by plants (Haber, 1905). During the following decades, between the 1930s and the 1960s, a second shift was produced when crop geneticists began to adapt cultivars to this new way of producing grains with high inputs (Castleberry et al, 1984). This combination of factors led to increase farm yields and grew the economic interest in understanding the crop response to inputs. This question has traditionally been addressed by running field trials that generate data, which is then analyzed and translated into recommendations for farmers.

Specifically, in corn (Zea Mays L.), this process became very active in the 1950s, 60s and 70s when the numbers of trials and publications increased (Heady and Pesek 1954; Heady et al. 1964, Heady, et al. 1964; Olson et al 1964, Hexem et al. 1976, Shrader et al 1966), opening a lively debate in major agricultural economics journals about the functional form of crop yield response functions. This conversation has continued over the last several decades (Swanson et al. 1973; Grimm et al. 1987; Frank et al. 1990; Berck and Helfand 1990; Paris 1992; Bullock and Bullock 1994; Chambers and Lichtenberg 1996; Llewelyn and Featherstone 1997; MaBL and Dwyer 1998; Anselin et al. 2006; Tembo et al. 2008; Tumusiime et al. 2011; Brorsen and Richter 2012).

In the first period, the conduction of the trials was done using labor-intensive techniques, with researchers marking small plots in the field, applying inputs and harvesting by hand or with small machines, and without the benefit of large-scale farm machinery (Bullock and Lowenberg-DeBoer 2007). This method was cost intensive, restricting the trials to small plots and mainly on experiment stations, where workers were available. Recommendations based on these small plots were extrapolated over large areas (Pan et al., 1997, Bullock, et al. 2002), ignoring variability and assuming average weather (Swinton and Lowenberg-DeBoer 1998). The results were useful at this initial stage only to provide a range of Nitrogen (N) rates that could be close to the EONR with a regional approach for a wide range of field characteristics. This regionals models were not useful to provide recommendations at a higher detailed level (field or site-specific). For that, more data, measuring more variables in different sites and weathers was necessary.

1

With the advent of precision agriculture technology starting in the 1990s, machines were able to control the applied N rate and measure different variables (like soil electro-conductivity, applied inputs rates or yield) site-specifically. This technological change shifted the demand for site specific N management recommendations (Bullock 2013). The previous regional recommendations were not sufficient to approach this new challenge, and new data needed to be generated that guided how different site-specific characteristics affected the EONR (Bullock et al 1998).

Fortunately, the same new technology could be used to generate agronomic experiments at low cost and with a high number of repetitions (Bullock et al 2002, Bullock and Lowendberg-DeBoer 2007). These advantages provided the groundwork for the following breakthrough in agronomic experimentation. The new concept is that the same farmers can run on-farm, large-scale field trials, over many fields to inexpensively gather large amounts of data in multiple field characteristic and weather (Bullock et al. 2009; Casanoves et al. 2007; Peralta, Cordoba, Costas, and Balzarini 2013). This new concept of on-farm Experimentation (OFE) is being implemented by collaborating researchers at the University of Illinois, University of Nebraska, Montana State University, Washington State University, and Louisiana State University, together with several international research partners, in an USDA-sponsored project called Data Intensive Farm Management project (DIFM). This work is originated in that project. The trials implemented by this working group have a "checkerboard" design and feature a large number of repetitions with plots small enough to explore similar site characteristics. The trials are also large enough to collect reliable data from the sensors. Researchers designed the trials for the farmers to conduct in their field, with minimal change in how they would regularly manage a production crop. Because experiments are run in an automatized way, it is possible to generate more data at lower cost than with the previously described labor-intensive plots from the past.

1.2 PURPOSE AND CONTRIBUTION

The project that originates this work has a working protocol, that states the instructions on how trials are designed, conducted and analyzed. The <u>objective</u> of this work is to explore the economic value of different strategies that could be used in that protocol, together with the environmental trade-offs from those field trials, and how that value can be optimized. With simulations we attempt to address: 1) Is the value of generating on-farm information in trials

likely to cover the cost of conducting such experiments; 2) What is the set of variables that, measured in the field and incorporated in the model, maximize profits for the farmer?; 3) What is the optimal number of years to run trials before moving to regular production using the results of the experimentation as management advice?; 4) What is the environmental benefit of these different strategies?; and 5) What insights can be used to detect the optimal number of trials in an ex-ante situation?

Answering these questions will provide insights that can be used to improve the protocol while helping the associated farmers of the project to make profitable use of the new concept. The results obtained in this work should not be taken as absolute values, rather as tendencies that will allow us to move on the path of discovering the best strategy.

1.3 WORK ORGANIZATION

This study is divided into six chapters. Following this introductory chapter, Chapter 2 provides a narrative review of previous studies that also used crop simulation in a spatially variable field. Chapter 3 is a review focused in providing the conceptual framework for the methods than will be used later in this study, paying special attention to models that explain and compare the value of information and technology. Chapter 4 describes the methods used in this research to create the field, simulate the crop under different situations and strategies, calculate a model using spatial techniques and finally analyze the results economically. Chapter 5 provides results from the simulations and discusses these results in detail. Finally, Chapter 6 summarizes this study's results and proposes avenues for future research.

CHAPTER 2: PREVIOUS STUDIES USING CROP SIMULATIONS IN SPATIALLY VARIABLE FIELDS

In Bakhsh et al (2001), the simulation model called Root Zone Water Quality Model (RZWQM), was used to evaluate the response of soils and crops to different N rates. They had a data set from 1996 to 1999 of a rotation of corn and soybeans trial data with nine plots where three different N-rates were applied. Tile flow, NO_3-N losses and yield was measured over the four years. The work showed three parts. In the first part, they calibrated RZWQM by adjusting some parameters to fit the model output with the real measurements of tile flow data from 1996, corn yield from 1996 and soybean yield from 1997. Then, in the second part, they evaluated the model using the data not used during the calibration. The model simulated annual tile flow adequately, by showing a difference of -8% between measured and simulated values. Similarly, yield for both crops showed a difference of less than 5% between measured and simulated values. Finally, in the third part, they use the model to simulate the effect of six different N rates, including the three tested. The results showed that RZWQM98 model could be used to assess the effects of N-applications rates on corn yields and NO_3-N losses in tile flow, and it may require further refinements in soybeans. Their article shares with the present work the idea of using crop simulation to understand the response of corn yield and N losses with different N rates. Their focus was on calibration and validation of the model, while in the present work no calibration was performed since we assumed that the model outputs are the real data. Instead, we focused on creating models that would transform that data into recommendations to farmers, which can then be evaluated in terms of farmer profits.

Thorp et al (2008) (30) developed a software called APOLLO that allows running DSSAT, a crop modeling software designed to model a crop in a uniform area, over a spatially variable area. The most significant incorporation is that it can condense, in a square grid, all the layers of information that the model needs and then run the model on each of the cells of the grid. In the second part of the work, they created a spatially variable field and described the process used to calibrate and validate the model using real data from that same field. Finally, in the third part they used the system to simulate the effects of climate change over the yield in the spatially variable field. For this they used two climate scenarios, one from the 1990s and the

4

other was a future prediction for 2040s. They conclude that climate change will decrease average yield in this situation and the spatial distribution of higher and lower-yielding areas.

The described software, APOLLO, follows a very similar procedure to the one we use in this work. We both created a grid condensing spatial information that is needed by the crop modeling software. The difference is that, instead of using a specially designed software to create the spatially variable grid and then run the crop modeling software, we performed this procedure using R (R Core Team 2018). Additionally, their crop modeling software was DSSAT and we used APSIM (Holzworth et al, 2014). However, the objectives of the two works are different, since they are predicting the effect of climate change over a field and we are creating trials and trying to fit models that help to provide advice based on those trials.

Although they described APOLLO in detail in the 2008 work, Thorp et al (2006) used it to jointly analyze the production function and the environmental externalities of different N rates in corn. In this work, they created a spatially variable field with 100 cells and, after calibration with real data, they run 13 N application rates over 37 growing seasons (based on 37 historic weather years) restarting the field every year to the assumed initial conditions. Their output included the yield of the crop and the amount of total N unused by the crop. They defined the profits maximizing rate for each cell as the one that maximized the mean profits of the 37 years. The environmental rate was determined to be the one that left less than 40 kg/ha of N in soil at harvest with a probability of 80%. In most of the cases, the profits maximizing rate was higher than the environmental rate. They calculated the opportunity cost of environmental protection as the difference in profits of decreasing the rates that were higher to the one that achieve the environmental goal. That opportunity cost was of $48.12 average for the whole field.

A similar approach to the previous article was followed in Miao et al (2006). In their work, they divided a field into four management zones based on cluster analysis of multiple soil layers. After calibrating and validating the model with real data, they used crop modeling with historical data of the previous 15 years to estimate the EONR, testing different strategies: ex-post estimation, an average of the EONR for each year and choosing the N rate the maximized the 15-year marginal net return. They mentioned that, although the first approach was the most profitable, it was not realistic since it is assuming full information. They concluded that choosing the N rate that maximized the 15-year marginal net return should be the method used by farmers

5

since it had higher long-term profits than the more straightforward method of averaging the yearly's EONR.

The two-last works, Thorp et (2006) and in Miao et al (2006) have much in common with the present work. The three works used crop modeling to estimate the response of corn to N in a spatially variable field. One important difference is that they used crop modeling to obtain the yield response curve and then EONR by running sequential rates for a same site of the field and calculating the optimal. In our work, each plot of our field had only one N rate each year, not a sequence. We pretended real trials over the field and crop simulation was used to start thinking about possible outcomes of those trials.

As presented in the previous review, there has been research using crop modeling software to simulate spatially variable fields with different objectives. To the best of our knowledge, no previous economic studies have used it to analyze the problem we will present here, combining this new concept of running whole field trials and using crop simulation to obtain the possible outcomes of those trials, compare strategies and finally optimize how those trials should be run to maximize farmer's profits.

CHAPTER 3: CONCEPTUAL FRAMEWORK

3.1 NATURE'S META-RESPONSE FUNCTION

To explain how to calculate the value of the new concept of OFE, first it is helpful to describe the hypothetic function of the response of Yield to inputs. Bullock and Bullock (2000) and Bullock et. al (2002) provided the "meta-response function" framework to discuss the concept of how crop yield (Y) responds to all the factors to which is responds. The function express the Y on a small and uniform site of a field as a function of a vector of "managed inputs" $\mathbf{x} = (x_1, \ldots, x_J)$ (decided by the farmer, such us hybrid, seed and fertilizer rates), a vector of unmanaged spatially dependent "field characteristics" $\mathbf{c} = (c_1, \ldots, c_K)$ (such as Organic Matter (OM), elevation, depth of the soil, soil N content) and a vector of stochastic time dependent variables called "weather" $\mathbf{z} = (z_1, \ldots, z_L)$ (mostly weather variables such as temperature, rain, radiation, first frost date, but also pest infestations and other factors):

$$y = f(\mathbf{x}, \mathbf{c}, \mathbf{z}) \quad (1)$$

$$c = f(\mathbf{p}) \quad (2)$$

In this work the concept of past variables $\mathbf{p} = (p_1, \ldots, p_m)$ is incorporated:

This vector is compounded by characteristics of a site measured in the previous season that affect the present conditions, such as Y_{t-1}, applied N_{t-1} rate and the ratio between them Y/N_{t-1} (where t-1 means previous year). These past variables do not affect Y directly, if not through an effect over $\mathbf{c}$. Nevertheless, they may worth being included in the models because with the new sensors technologies, they have the advantage of being easily measurable at a low cost and being highly available on most farms. One goal of this work is to explore how using $\mathbf{p}$ in the function would be a promising cost-efficient method to predict the Y response function. For example, initial N in the soil is a $\mathbf{c}$ variable that can be obtained by doing soil sampling (SS), incurring a certain cost. Instead of using that variable, we can use Y/N_{t-1}, expecting that when it is high, the previous crop had used most of the applied N and the residual N in the soil would be low and vice versa. That way, an important variable of the function, like initial N, can be predicted using a low-cost variable.

Based on the knowledge of the farmer about the three different vectors, it is possible to estimate an expected response function. That function can be analyzed and optimized to generate

7

a field's application map that maximizes the expected profits for the farmer based on the available information. When more information is available, the farmer's decision should be closer to the true EONR.

3.2 CALCULATING THE VALUE OF TECHNOLOGY AND INFORMATION

In the first stages of precision agriculture, many studies were performed on the economics of the new technology. Unfortunately, these studies used a wide range of assumptions and methods (Lowenberg-De Boer 2003). Some of them omitted significant cost, made different assumptions, or compared profits under different situations. A common mistake had been using rates obtained without trial data for Uniform Rate Application (URA) versus rates obtained by analyzing trial data for Variable Rate Applications (VRA), confusing the value of technology with the value of information (Bullock et al. 2009). A Purdue University review from 2000 showed that the number of economist co-authoring articles was increasing between 1991 and 1999, and several authors outlined methods for obtaining reliable economic estimates of the new technology (Swinton and Lowenberg-DeBoer 1998, Bullock and Bullock 1998).

Bullock et al. (2009) used a methodology to compare the value of different strategies. They compared three factors that affect the decisions of the farmer:

- Y response to N knowledge: the knowledge of the exact response function of Y to N rate (ex-ante, ex-post). "Ex -ante" means that decisions are made before the growing season, thus the weather is unknown and the farmer does not know the exact yield curve, thus he must use some historic average. "Ex-post" means that different rates are tested during the growing season, the exact response curve is known and the optimal N rate selected for each situation. In "Ex-post" scenario, decisions are optimized for the growing season.
- Information (I): the availability of site-specific information for the field (yes, no). When more significant spatial variables are incorporate as predictors to a model, the capacity to explain and predict the response is usually increased. In their work, no information meant that the farmer had no knowledge of site-specific characteristics and, in that situation, he had no incentive to use VRA and he would use URA and a regional model to predict the EONR.
- Technology (T): weather the farmer is using uniform rate application (URA) or has access to Variable Rate Technology (VRT).

8

By combining these three factors, eight structures were created combining information and technology levels.

It can be easily seen how technology can increase profits, by comparing the achieved profits with and without the technology. It is more complex to see how information affects profits. According to Lowenberg-DeBoer (1998), information affects profits if it changes decisions. If certain decisions are more profitable than the ones that would have been made without the information, the increase in profit is due to the information. Following that reasoning, a farmer will choose the N rate that maximizes his expected profits on each of the eight structures. Then, by comparing the change in actual profits, the marginal value of each structure can be determined.

The economic analysis of this work built over that approach, with slight adaptations. The first adaptation is that EONR was calculated only in ex-ante conditions. This is because the interest of the work was analyzing real strategies under partial-information (unknown weather) and thus, the ex-post EONR is not meaningful for this goal.

The second adaptation is how information is treated. At present, most of the modern machines used in the U.S. Corn Belt come with sensors to measure different variables. Since the price of these sensors has been decreasing over time, nowadays most of the base models of machinery include them. Moreover, that trend is expected to increase in the future. For that reason, these sensors are assumed a sunk cost, and the information they provide is assumed to be free. Researchers in the project, in constant contact with farmers, state that it is common to observe that modern farmers have accumulated over the years many as-applied maps (that is a map with points where applied rates of the planter or sprayer are recorded), yield maps (that is a map with points were yield at harvesting is recorded) and elevation maps from their fields over time, and they are willing to transform it in information that can be useful. For that reason, instead of having two levels of information (yes or no), three levels were created. One is the "no" information level and includes the situation where a farmer does not have -or if he/she has does it is not used for N management purposes- site-specific information and is not running trials. The "low" information level is the situation where a farmer run trial/s and has free site-specific information provided by sensors at no cost or insignificant cost. Finally, the third is the "high" information level and it is the situation where a farmer has information from trials together with

free site-specific information and also gathers soil sampling information every year for initial N and every four years for OM.

3.3 MAXIMUM RETURN TO N (MRTN)

In 2004, University soil fertility researchers from different States in the corn belt (Illinois, Iowa, Michigan, Minnesota, Ohio and Wisconsin) with the aim of unifying methods started a program oriented to provide regional N rate guidelines (Sawyer et al. 2006). In this program, they collected information from trials across the region, analyzed the data to generate response models and create, as a final output a N-calculator that is available to the public in their website (http://cnrc.agron.iastate.edu/).

They way how N-calculator works is allowing the user to select an area within the 6 states, together with management information (rotation, type of fertilizer) and prices of N and corn. Then the platform provides the result of the model in those conditions, and the suggested rate to apply to get the Maximum Return to N application (MRTN rate). For the region where the field is located, the MRTN rate is 224 kg/ha of N.

CHAPTER 4: METHODS

4.1 CONCEPTUAL STEPS OVERVIEW

The present work is compound by two parts. First the data generation process and then the analysis process. Figure 1 shows the workflow followed during the data generation process. First, a spatially variable field was created. Over that field, one year of a regular crop using MRTN rate was simulated to set the initial conditions. That simulation year is not part of the observations. The research process starts with the *Trials Stage* (TS), where completely randomized trials were simulated each year. After TS, the trial information was analyzed, and an econometric regression model was estimated. The *Production Stage* (PS) is considered the stage when the field is used for the regular production of a crop using the EONR obtained from the trial data.

A ten years-long simulation for the whole field (256 cells x 10 years = 2560 observations), composed by TS and PS is called a *set* $(z,L,S^{I,T})$. Different sets were created by combining the following factors:

- Weather scenarios (z = 1980, 1985, 1990, 1995, 2000): They were created using historic weather and named based on the first year of the sequence. That means that z = 1980 uses the weather from 1980 to 1989. Consecutive years of the historic weather were used, to capture real patterns in the weather.
- Number of trials (L = 0 to 5): being 0 when no trials were run and L from 1 to 5 the number of years when a trial was run over the field.
- Information (I = No, Low, High): No is the situation where no trials were run; Low is where trials were run and "free" site-specific variables are collected. High is where trials were run, "free" site specific information is collected and also soil sampling information of initial Nitrogen (N^{Apr}) and Organic Matter (OM^{Apr}).
- Technology (T = URT, VRT) was the technology used during PS, being URT when the farmer was constrained to select only one rate that maximizes expected profits in the whole field VRT when the farmer had the ability to select the rate that maximizes expected profits in each site of the field. It is important to note that during TS the farmer will always use VRT, because changing the rates that is needed to apply the different treatments in the trial.

11

Strategies ($S^{I,T}$) was the word used to describe the possible combinations of Technology and Information during this work. Since comparisons were done usually by z and at L^{OPT}, this will simplify the notation and avoid repeating the first two superscripts.

In this work, it was assumed that a farmed that did not run trials will not have incentive to use VRA, and will use only URA of MRTN rate. That way, five S ($S^{no,URA}$, $S^{low,URA}$, $S^{low,VRA}$, $S^{high,URA}$, $S^{high,VRA}$) were obtained in this work, tested over different L and z.

Having explained what a set and S are, it is possible to account the total number of sets simulated in this work. The sets are combinations of z,L and S, but not all combinations are possible. The number of sets that do not involve trials are 5 and they are the combination of $S^{no,URA}$ with L = 0 (if no trials are run, L can only be 0). The number of sets that involves trials are 100, and they are the combination of the five z (1980,…,2000), with the five L (1,…,5), with the remaining S that includes trial information ($S^{low,URA}$, $S^{low,VRA}$, $S^{high,URA}$, $S^{high,VRA}$). Adding them, the total of 105 sets simulated in this work are accounted. Considering that each set is 10 years long and that the whole field has 256 cells, that is a total of 268,800 simulations.

Figure 1: Data generation process workflow
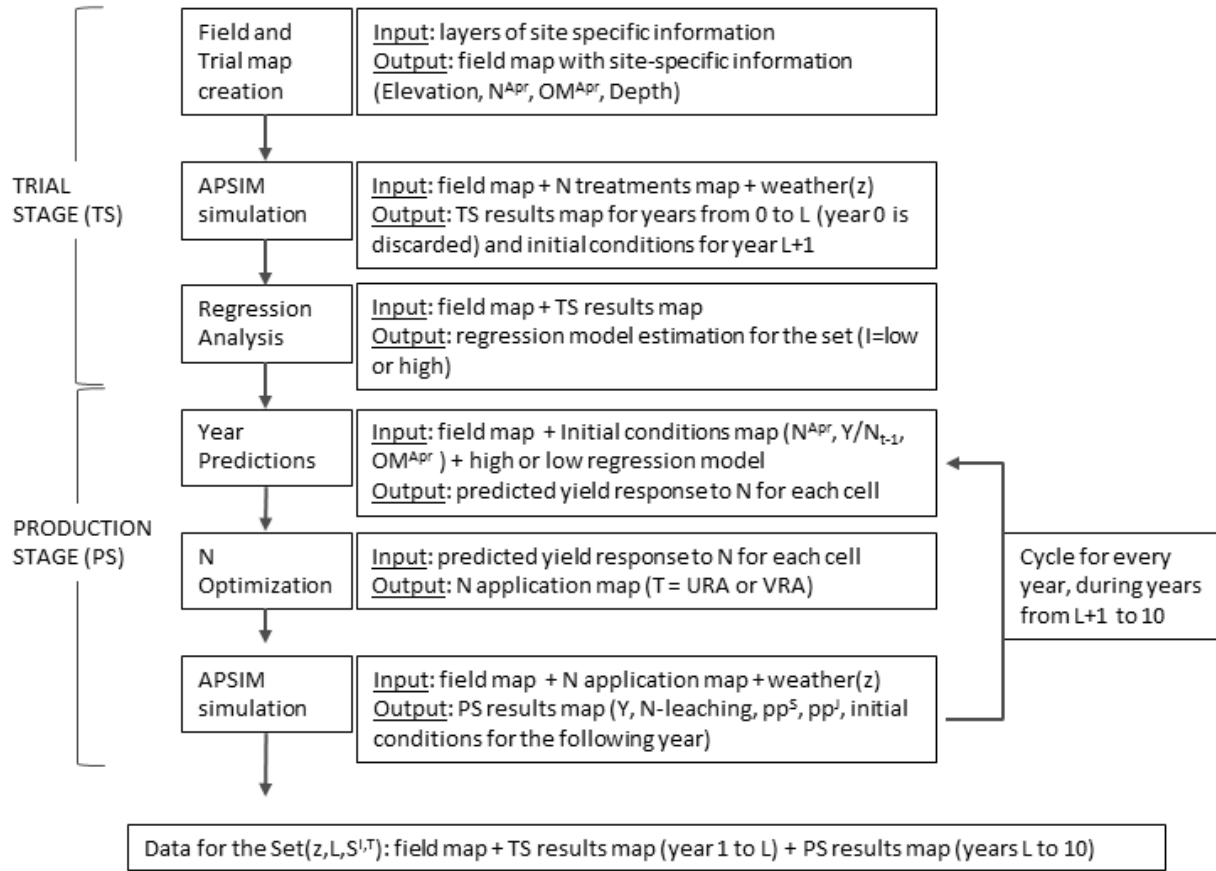
After generating the data, the work continued with the analysis process shown in figure 2. First, economic variables were computed. The input was the APSIM results map. This map was condensed into total field values by year (i.e. the N and Y by cell was added for the whole field to obtain the total N applied and the total grain harvested) and the costs and revenues were computed by year.

The trajectory of the Present Value (PV) of each set by increasing L was graph and analyzed to answer question 1 from section 1.2. Using the PV at the L^{OPT}, the value of the different combinations of I and T was obtained, answering question 2 and 3. Then, Environmental Variables trajectory was analyzed to answer question 4. Finally, two especial analyses were performed to address question 5, one to understand underlying causes of the results and the other one to find insights that could help to predict the L^{OPT} in ex-ante conditions.

Figure 2: Data analysis process
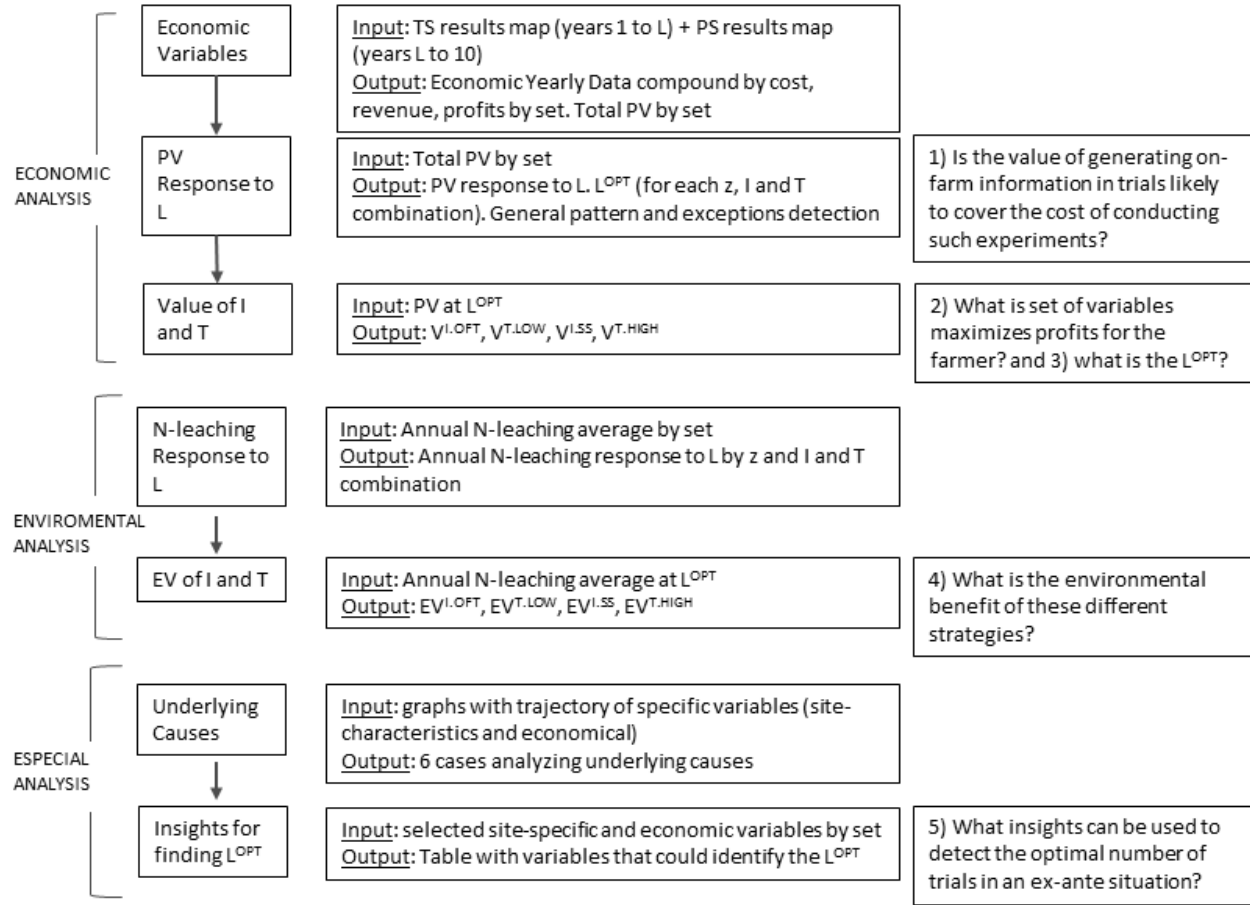
4.2 FIELD CREATION

4.2.1 INITIAL CONDITIONS

To capture the essence of spatial experimentation and management, a farm field with spatially heterogeneous field characteristics was modeled. A real 32–ha field owned and managed by a farmer near Effingham, Illinois (39.1 N latitude and 88.7 W longitude) was used as starting point.

Spatial site-specific information about elevation, information from soil sampling, and the soil survey report (SSURGO 2018) were collected. All the variables were edited, following different procedures. These way, we created a new research field, different than the original, that could lead us to explore how variability affects the results. The final product is a square grid (figure 3) with one-ha cells (c = 1,…,32). Each cell had unique site-specific characteristics obtained by condensing and editing the different layers of site-characteristics information with the following procedures:

14

- Soils: because the soil survey report indicated that 85% of the field is classified under the same soil type, for simplicity we categorized the whole field under that soil type. This soil is named Cisne silt loam, with 0 to 2 percent slopes. The typical profile is Ap-E-Bt1-2Bt2-2C. These soil characteristics were reproduced in APSIM following the methodology described in Archontoulis et al. (2014).

- Organic Matter in April (OM^{Apr}): OM was sampled in April 2016 at 32 points, each at the center of 1-ha grid cell. Since OM was not highly variable over the field, with a mean of 2.42% and standard deviation of 0.15, the residuals between each measurement and the mean were multiplied by 1.8 to increase the variation. The final OM^{Apr} has a mean of 2.42% and a standard deviation of 0.27. The purpose of this was to explore more variability in the field that the one present in 2016.

- Nitrogen in April (N^{Apr}): is the sum of the quantities N available as N-NO$_3$ and N-NH$_4$ in April, the day before planting. It was estimated by conducting a linear relation with OM. Also, a spatially autocorrelated error was added to the result of the function using Gaussian simulation, according to the following autoregressive process:

$$N.initial = 25 * OM - 38.52 + Q \hspace{3cm} (1)$$

$$Q = \rho_q WQ + error$$

$$error \sim N(0, \sigma_q^2),$$

where ρ_q is the spatial autoregressive parameter, W is the spatial weight matrix, and μ_q and σ_q^2 are the expected value and variance of Q. The W is a matrix whose (i,j) elements are $e^{-0.1*d_{ij}}$, where $d_{i,j}$ is the distance between the centroid of the *i*-th *and j*-th square of the grid.

With this procedure a spatially autocorrelated N^{Apr} was obtained for each cell of the field. N^{Apr} values varied among cells between 1 kg/ha and 37 kg/ha.

- Elevation: using high-density point information data from the tractor GPS system, a point-in-polygon operation was performed, obtaining the median of the elevation points for each square of the grid.

- Depth: The original thickness of the soil units' last layers varied between 153 cm and 196 cm. Spatial variation of the total soil depth was generated by varying only the thickness

of the last layer (upper layers were not affected by this procedure). In the same way than with N^{Apr}, Depth was calculated as a linear function of Elevation, and an autocorrelated errors were added:

$$Depth = 3.183\ Elev - 1806.46 + Q \hspace{3cm} (2)$$

$$Q = \rho_q WQ + error$$

With this procedure, a spatially autocorrelated depth was obtained for each cell of the field, which varied from 153 cm to 200 cm among cells.

As explained before, a set is a 10-year simulation of the whole field. Every set starts with the field and the initial conditions described above. Then, a regular corn crop was simulated using a nitrogen fertilizer application at the MRTN rate. These simulations of a regular corn crop were considered to be in year 0 and they were discarded. Since the initial conditions have an impact on the result of the trial, the goal of simulating a regular corn before the TS was to expose those initial conditions to a simulation with a random weather. If this were not done, the first trial year would always have had the same initial conditions, which would impose the same effect over the L = 1 strategy in all z scenarios. For example, if the initial conditions have a high N^{Apr}, then response to N would be low for L=1 and random for higher values of L, which is not realistic.

4.2.2 TRIAL DESIGN

Over the previously described field, a trial was created following the protocol used in the DIFM project. A planter machine width of 12 meters was assumed. A border of 24 meters was created on the edges of the field and considered an area excluded for experimentation. The reason is that, in real trials, the edges of the field are exposed to uncontrolled factors that decreased the quality of the data, like winds, heavy machinery transit or different soil quality. Even though these factors will not be incorporated in the crop modeling simulations, the border was included because it will change the relative area of the trial and, since the border will receive the target rate instead of the different treatments, it will have an impact on the economic result of the strategy.

As shown in figure 3, inside the border, a grid of 110 plots was created. Each plot had 24 m width and 85 m length. When overlapping the initial grid of varying **c** characteristics, with 32

16

squares of 1 ha, with the trial design with 110 plots and a border, 256 cells are gotten. Within these cells, rather the site-characteristics or the received treatment was different. For that reason, each of these cells was unique and all APSIM simulations were run for each cell. Then, results were aggregated, to a dimension that represents what we would obtain in a real situation. For example, if at the time of analyzing the data from the trials, a plot had three cells, all needed variables were aggregated weighing by area and that that was the data used for the regression analysis.

With the aim of using the same naming convention throughout the present discussion, each of the 32 squares in the initial grid was called a "*square*". Each of the 110 trial plots that was created over the field was called a "*plot*". Each of the 256 cells with varying dimensions that have a unique combination of site-characteristics with treatment is called a "*cell*".

Figure 3: Naming convention for field layers and resolution
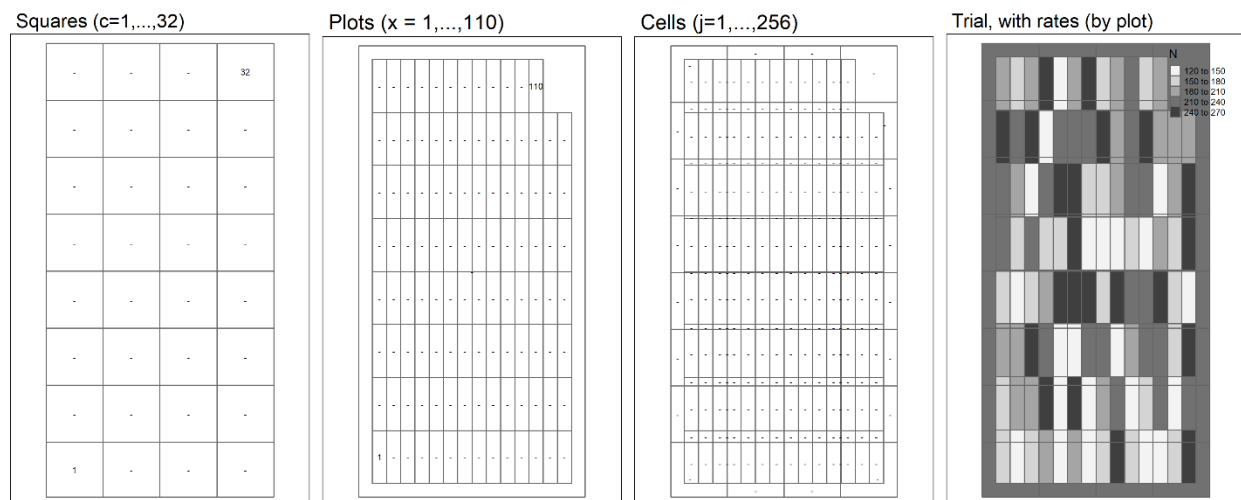
4.2.3 DYNAMIC TREATMENTS

Farmers run annual field trials with a completely randomized design. Each year, five nitrogen fertilizer rates were randomized on the 110 plots. That is, each one of the five N Rates were applied on 22 of the 110 plots. The treatments (T) were selected dynamically based on the following process:

Each year there was a central target rate (TR) and the other 4 rates were calculated based on this central rate according to: T1= TR – 90 kg/ha, T2 = TR – 60 kg/ha, T3 = TR – 30 kg/ha, T4 = TR, T5 = TR + 30 kg/ha. Areas along the field perimeter were placed in a "buffer zone," not included in the experiment, and assumed fertilized at rate TR. The first trial year, the TR was

the MRTN rate. In subsequent years, the TR was the average among the MRTN rate and the rates that would have maximized profits in each of the previous trial years.

4.2.4 OTHER MANAGEMENT PRACTICES

All other management practices that are not related with N rate were held constant in all simulations. Soil sampling was performed on April 29th (as explained later, this information could be used or not, depending on the strategy). The field was tilled every year on April 30th of every year, at a depth of 11 cm, incorporating 40% of the surface residues. The planting date was May 1st of every year. The hybrid was a generic APSIM hybrid with called A110, with 110 days maturity. The plant population was 8 plants/m^2 and the row spacing was 76 cm. No tiling was included in the field.

4.3 ON-FARM YIELD CURVE ESTIMATION

4.3.1 SOIL SAMPLING SIMULATION

To make results economically realistic, soil sampling variables needed to be transformed. High-resolution N^{Apr} and OM^{Apr} data were provided for each of the 256 cells, by APSIM every April. Since soil sampling at that resolution would be economically prohibitive, two new variables were created. They were the spatial and temporal aggregation of the high-resolution N^{Apr} and OM^{Apr} data. Although this process decreased the precision of the data, it is imitating what could be obtained in a real situation.

One of the new variables was the *soil N in April obtained by soil sampling* ($N^{Apr.ss}$). This variable was the area-weighted average of N^{Apr} (whose resolution is by cell) with the following soil sampling strategy:

- Year 1: by square. Since the field previously had a uniform-rate crop, variation will depend on site-specific characteristics that varied by square.
- Year 2 to L +1: by plot. During trial years, since treatments were applied by plot, the soil sampling was performed at that same scale. Then in the first year after TS (year = L + 1), since the field still would have high variability in N^{Apr} because it had different rates in the previous year, $N^{Apr.ss}$ is still performed by plot (and the border area was characterized by a single soil sample).

18

- Year L+2 to 10: by square. Two years after the TS, N^{Apr} variation over the field is smoother, and soil sampling is performed by square.

The other new variable was the *soil OM in April obtained by soil sampling* (OM$^{Apr.ss}$). This variable was the area-weighted average of OM^{Apr} simulating a soil sampling strategy performed by square every four years (that is in April of year 1, 5 and 9). That means that for one of the squares OM$^{Apr.ss}$ had the same value from year 1 to 4, then a new value from year 5 to 8, and finally another new value from year 9 to 10.

4.3.2 MODEL ESTIMATION

In section 3.2 and 4.1, the concept of how Information was treated in this work was introduced. The level of information of the set will impact the variables included as regressors in the model estimation. There are three levels of information (no, low, high). If the information level was no, that means that the farmer did not conduct trials, thus no regression model was estimated.

If the information level was low, the model included variables that are considered "free" nowadays. They are the as-applied N, elevation, Y/N_{t-1}. The N rate is automatically measured by sensors in the fertilizer applicator that automatically records the rate. The Elevation is recorded by the guiding system of the tractor. Y is measure by the yield monitor in the harvesting machine and then transformed to the Y/N ratio. The Low information model was:

$$Y_{it} = \beta_0 + \beta_N * N_{t,i} + \beta_{N^2} * N_{t,i}^2 + \alpha_2 * Year2 + \alpha_2 * Year2 * N_{t,i} + \tag{3}$$

$$\alpha_2 * Year2 * N_{t,i}^2 + \alpha_3 * Year3 + \alpha_3 * Year3 * N_{t,i} + \alpha_3 * Year3 * N_{t,i}^2 +$$
$$\alpha_4 * Year4 + \alpha_4 * Year4 * N_{t,i} + \alpha_2 * Year4 * N_{t,i}^2 + \alpha_5 * Year5 + \alpha_5 *$$
$$Year5 * N_{t,i} + \alpha_5 * Year5 * N_{t,i}^2 +$$

$$\beta_E * E_i + \beta_{EN} * E_i * N_{t,i} + \beta_{EN^2} * E_i * N_{t,i}^2 + \beta_{Y/N} * Y/N_{t-1,i} + \beta_{Y.N * N}$$
$$* Y/N_{t-1,i} * N_{t,i} + \beta_{Y/N*N^2} * Y/N_{t-1,i} * N_i^2 + \varepsilon_{t,i}$$

with $\varepsilon_{t,i} = \lambda W \varepsilon_{t,i} + u_{t,i}$ and $u \sim N(0, \sigma^2 In)$

Where,

t is the trial year and goes from 1 to 5, i is the plot and goes from 1 to 110,

19

β_0 is the parameter for the intercept,

$Year2$ to $Year5$ are dummies variables for the respective year. When combine with the respective parameter α_2 to α_5 they constituted the year effect, representing the change in the quadratic response relative to year 1 (omitted category). They estimated unobserved characteristics that are specific to a particular year,

$N_{t,i}$ is the N rate, E_i is the Elevation, $Y/N_{t-1,i}$ is the Y/N ratio from the previous year,

If the information level was high, the model included all the low information variables, but also incorporated the soil sampling information. The High information model was:

$$Y_{it} = \beta_0 + \beta_N * N_{t,i}^T + \beta_{N^2} * N_{t,i}^{T\,2} + \alpha_2 * Year2 + \alpha_2 * Year2 * N_{t,i}^T + \qquad (4)$$

$$\alpha_2 * Year2 * N_{t,i}^{T\,2} + \alpha_3 * Year3 + \alpha_3 * Year3 * N_{t,i}^T + \alpha_3 * Year3 * N_{t,i}^{T\,2} +$$

$$\alpha_4 * Year4 + \alpha_4 * Year4 * N_{t,i}^T + \alpha_2 * Year4 * N_{t,i}^{T\,2} + \alpha_5 * Year5 + \alpha_5 *$$

$$Year5 * N_{t,i}^T + \alpha_5 * Year5 * N_{t,i}^{T\,2} +$$

$$\beta_E * E_i + \beta_{EN} * E_i * N_{t,i}^T + \beta_{EN^2} * E_i * N_{t,i}^{T\,2} + \beta_{Y/N} * Y.N_{t-1,i} +$$

$$\beta_{Y/N * N} * Y/N_{t-1,i} * N_{t,i}^T + \beta_{Y/N*N^2} * Y/N_{t-1,i} * N_{t,i}^T + \beta_{OM} * OM_{t,i}^{Apr.ss} +$$

$$\beta_{OMN} * OM_{t,i}^{Apr.ss} * N_{t,i}^T + \beta_{OMN^2} * OM_{t,i}^{Apr.ss} * N_{t,i}^{T\,2} + \varepsilon_{t,i}$$

with $\varepsilon_{t,i} = \lambda W\varepsilon_{t,i} + u_{t,i}$ and $u \sim N(0, \sigma^2 In)$

Where,

$N_{t,i}^T$ (Total Nitrogen), is the sum of $N_{t,i}^{Apr.ss}$ and $N_{t,i}$ rate,

$OM_{t,i}^{Apr.ss}$ is the OM obtained by soil sampling.

In statistics, Ordinary Least Squares (OLS) is an estimation method for the unknown parameters that assumes that errors in the dependent variable are uncorrelated with the independent variable(s). When this is not the case, OLS would not provide unbiased model estimates. For that reason, in this work, statistic procedures that allow spatial autocorrelation, were used to estimate the unknown parameters.

In spatial statistics there are different options for including spatial autocorrelation in a model. Given the characteristics of the variables in this work ($N_{t,i}$, E_i, $OM_{t,i}^{Apr.ss}$, $N_{t,i}^T$, $Y/N_{t-1,i}$), no spillover could exist between the different plots. The treatments were applied only in the target area, affecting the Yield of that site and not the neighbors. Moreover, a change in the regressors would not affect the Yield of a neighbor indirectly (no spillover effect). For that reason, the only model for spatial autocorrelation that could be used is the **Spatial Error Model (SER).** When only one year of data was available (L=1), the "errorsalm" function from the R package "spdep" (Bivand et al. 2013, Bivand and Piras 2015) was used. If more than one year of data was available (L > 1) the Spatial Panel Error model from the R package "splm" was used (Giovanni and Piras 2012).

4.3.3 SOURCES OF ERROR FOR THE MODEL

In this work, data was simulated using APSIM based on all the provided inputs. Then, a regression model was adjusted over that data using less information, or with different spatial and temporal resolution. In consequence, the regression did not explain all the variability, having a random error - $\varepsilon_{t,i}$ in (1) and (2). In this section, the sources of that random error are explained in detail:

-   Omitted variables: Depth was one of the inputs in the APSIM simulations and it affects Yield, especially in dry years, since it changes the amount of water that the soil can retain. This variable was not included in the regressions directly, it was included indirectly since it is correlated with Elevation. The two other omitted variables were $N^{Apr.ss}$ and $OM^{Apr.ss}$ in the low information model. The portion of the variability explained by the omitted variables would be part of the error in the model.
-   Soil sampling representation: since soil sampling is costly, in the high information model, N^{Apr} and OM were spatially and temporally aggregated based on the procedure explained in 4.3.1. This decreased precision of the data, increasing the error of the model.
-   Spatial resolution: The field had 256 cells and APSIM is run every time for each cell. One plot can be composed by more than one cell. For the regression model estimation, variables were aggregated doing and area-weighted averaged to obtain one observation per plot (total of 110 observations). All the variables in the model were affected by this process Y_{it}, N_{it}, E_i, $OM_{t,i}^{Apr.ss}$, $N_{t,i}^T$, $Y/N_{t-1,i}$).

21

In the case of $N_{t,i}^T$ and $OM_{t,i}^{Apr.ss}$ they were aggregated twice, first according to the soil sampling resolution and then, if a plot had more than one value, it was again aggregated by plot.

- Functional form: APSIM is not an equation, it is a crop modeling software with modules that interact which each other. In this work, a quadratic functional form was used and it has not necessarily the exact same shape than the APSIM output, increasing the error term.
- Weather not captured by year fixed effect: To make the model applicable in a wider range of situations, the year effect was incorporated only affecting the quadratic response of Y to N and the intercept, not the other regressors. If it also interacted with other variables, the interaction will be part of the error term.

4.4 EONR PREDICTION

4.4.1 WEATHER WEIGHTED PROFIT FUNCTION

During PS, every April, yield predictions for each cell for each year were obtained using the corresponding regression model, the initial conditions of the cell in April for that year and an increasing sequence of N rates. As a precaution, the included sequence ranged from the lowest tested rate in the trial to the highest, to avoid extrapolations outside the tested rates.

Some considerations were made regarding the year dummy variables. The model for L > 1 had a fixed year effect variable interacting with the intercept, the linear and the quadratic response to N (or N^T) in the high information model. These year fixed effects are for past years and, at this time of the process, we need to predict the EONR for the future year, in ex-ante conditions, with uncertainty in the weather. To do this, the yield for each N in the sequence rates was calculated using each year fixed effect. Then, Partial Profits were calculated considering the price of corn (P^C= 0.157 \$/kg = 4 \$/bu) and the cost of Nitrogen (P^N = 0.881 \$/kg = 0.4 \$/lb). All other costs are assumed to be the same for all rates and not considered at this part of the work since they will not affect the optimization process. For example: if L = 3, a dummy variable for year 2 and year 3 was added to the sequence of N rates, and three response curves of profits to N were created for each cell.

The profit curves for each year were condensed in one **weather weighted profit function**. For this, a probability of occurrence was assigned to each of them. This probability

was calculated using the historic weather and assuming a normal distribution of it. Two weather variables were used: Season precipitation (pp^S) and July precipitation (pp^J):

- pp_t^S: is the amount of precipitation from sowing to harvesting during year *t*. pp^S affects two processes. On one hand, if the crop is limited by water, an increase in pp^S would produce a higher response of the yield to N, because it will increase the growing rate of the plants and with that the uptake of N. On the other hand, an increase in pp^S is associated with an increase of leaching, especially if the precipitations are concentrated in the stages were the growing rates (and thus the N uptake) are low.

- pp_t^J: is the amount of precipitation in July of year *t*. July is the month of the year where a crop planted at the beginning of May will be flowering. The -15 and +15 days before and after flowering are the Critical Period of the crop, when the crop generates the potential grains and thus the availability of resources produce a great impact on yield. As a result, high values of pp^J are associated with a higher capacity of the crop to uptake N and generate in Y. In contrast, water limitations in this period will reduce the Y of the crop and the demand of N.

Is important to consider both variables, because they could have different effects on the crop-soil system. For example, it is expected that a high pp^S in combination with a low pp^J will produce a high N-leaching without increase in Y. This is because the precipitations will be concentrated in a period of the season when the crop is not growing fast and up-taking N, thus the N in the soil will be transported below the roots depth. If a normal pp^S is combined with a high pp^J it is expected to have a high response of Y to N, and not much over the N-leaching, since the crop is receiving water in the moment when it is at full capacity to uptake the nutrient from the soil and generate grains.

For that reason, both variables were combined to determine the probability of each year, doing the following procedure:

With Season precipitation		With July precipitation	
$$Z_t^S = \frac{pp_t^S - \mu_{pp_t^S}}{\sigma_{pp_t^S}}$$	(5)	$$Z_t^J = \frac{pp_t^J - \mu_{pp_t^J}}{\sigma_{pp_t^S}}$$	(7)

$$PZ_t^S = P\left(-abs(Z_t^S)\right) \qquad (6) \qquad\qquad PZ_t^J = P\left(-abs(Z_t^J)\right) \qquad (8)$$

Combined probability of both weather variables:

$$WP_t = \frac{PZ_t^S\, PZ_t^J}{\sum_{t=1}^{t=L}\left(PZ_t^S\, PZ_t^J\right)} \qquad (9)$$

Where μ is the mean and σ is the standard deviation of the indicated variable. Then, Z is the distance between each of the observed values and the respective historic mean in units of standard deviation.

$P\left(-abs(Z)\right)$ is the accumulated probability of the negative of the absolute value of the Z score (left side of the distribution), assuming that $Z \sim N(0, 1)$

With this method, the Weather Probability (**WP$_t$**) of each profit function was calculated considering how usual was the combined weather of both variables. This allowed to decrease the weight of trials run in years with a weather that is not expected to happen in the future, and increase the weight of those whose weather is more likely to happen.

To explain this, the example in table 1 is provided. With z = 1990, 5 trials were conducted (L = 5). The historic weather from 1979 to 2015 has a $\mu_{Pp_t^S} = 454$, $\sigma_{Pp_t^S} = 153$, $\mu_{Pp_t^J} = 105$ and $\sigma_{Pp_t^J} = 62$. In this example, year 4 was an extremely wet year, both for the total amount during the Season and during July. If we calculate the profits as a regular average of the 5 productions functions, we would assign year 4 a probability of 20%. With this methodology we will assign it a probability of 0.1%. Year 3 had a Pp_t^S very close to the mean and thus, highly probable. Nevertheless, Pp_t^J was extremely high, with a very low probability. This last characteristic decreased the final probability of year 3 to 0.5%. The weather weighted profits function will be mainly a combination of year 2 (61.9%), year 1 (21.8%) and year 5 (15.5%).

Year	pp_t^S	Z_t^S	PZ_t^S	pp_t^J	Z_t^J	PZ_t^J	WP_t
1	635	1.183	0.118	127	0.354	0.361	0.218
2	387	-0.437	0.330	84	-0.338	0.367	0.619
3	483	0.189	0.424	279	2.806	0.002	0.005
4	762	2.013	0.022	244	2.241	0.012	0.001
5	297	-1.026	0.152	53	-0.838	0.200	0.155

4.4.2 RESOLUTION AND TECHNOLOGY

In section 4.2 and Figure 3, the naming convention of square, plot and cell was stated. The model was generated using as data the trial information, where each of the 110 plots was an observation. The reason why data was aggregated to one observation per plot is that, under a real situation, it takes time and distance to the spraying machine to change from a rate to another, especially when rates are separated one from another like in a trial. Also, the yield monitor has a lag between it finishes processing one plot and the harvesting machine has enough flow of grain from the following plot, to provide a quality measure of grain.

During PS, the change in rates over the field is expected to be smoother than over the trial and it was assumed that the sprayer could apply the target rate and that the yield monitor was able to provide quality measures of yield for each of the 256 cells. For that reason, even though the model was generated at the plot level -i.e. using plots as observations-, the predictions were done at the cell level.

After calculating the weather weighted profits function for each of the squares there are two possible paths to proceed:

If the farmer was going to use URA during PS, he had the restriction of using only one rate in the whole field. For that, all the weather weighted profits function for each cell were combined in only one for the whole field. This was done by adding the gross profits of each cell (considering the area) for each N rate. Then, the N rate that maximized the profits for the whole field was selected ($EONR_{UR,t}$). This is the rate that will be applied the following year uniformly over the field.

On the other hand, if the farmer is using VRA during PS, the weather weighted profits of each cell for each N rate was calculated. Then, the N rate that maximized profits was selected ($EONR_{VR,t,j}$). This was the rate that will be applied the next year on each cell, using VRA.

4.4.3 SOURCES OF ERROR FOR THE PREDICTIONS

The predictions of yield and profits for a future year were made using the model that has a random error term $\varepsilon_{t,i}$ whose composition was explained in 4.3.3. Since the predictions are made for future conditions, instead of past, new sources of error are added:

- Different weather: predictions are obtained for a future year, with unknown weather. The weather weighted profit function created with the historic weather could be different than the one that will be experienced during the next growing season.

- Extrapolation in initial conditions: the initial conditions in April can have a different combination of values in the different variables of the model than the ones explored during the trial years. Thus, the predictions would extrapolate values outside the evaluated range that could lead to error in the predictions. For example, the model could be created with a range of variation in Y/N$_{t-1}$ and applied afterwards in a different range.

- Spatial aggregation of N^{Apr}: After TS, N^{Apr} is aggregated. The first year, the aggregation is at a plot level and then at a 1 ha square grid level. This reduce the precision of the model.

- Evaluated rates: each trial had five treatments rates, with rates separated by 30 kg/ha. Prediction of the Y response to N were calculated using a N sequence that goes from the maximum N rate used in the trial to the minimum N rate use in the trial, by 10 kg/ha . In consequence, there is interpolation between the rates used for generating the model and the rates used for the predictions.

4.5 ECONOMIC ANALYSIS

For each set(z, L, S^{I,T}), all the cost related with N management was tracked and considered in the economic analysis. During TS, the cost of the farmer's extra effort and time that takes to learn and implement the experiment on year t was called C_t^F. In the DIFM project, new farmers that are starting to participate in the project for the first time, are invited to a two-hour meeting where they are informed about the basic experimental procedures. Including the

26

travel and the time, an opportunity cost of $200 was determined for attending this meeting. Also, researchers of the project have noticed that there is a learning curve of farmers, where the first year they have frequent questions related with the conduction of the experiments, and in subsequent years consultation becomes less necessary. For that reason, a lump sum of $500 was considered the first year as a compensation for the communication effort and time of the farmer. This sum was decreased to $400 for year two, $200 for year 3 and $100 for the remaining years of experiments. Adding both items, C_1^F=$700, C_2^F=$400, C_3^F=$200 and $C_{t>3}^F$=$100 for the remaining years

C_t^A was the total cost of applying N on the whole field during year t. The price charged by fertilizer applications service providers vary in the U.S. Corn Belt. Plastina, Johanns, and Erwin (2016) reported an average charge of 16.43 $/ha for URA of liquid fertilizer from a survey in Iowa. For those same services, Miller (2013) reported an average charge of 15.17 $/ha from a survey of Indiana, Stein (2014) reported an average charge of 18.78 $/ha by custom applicators in Michigan, and Halich (2016) reported an average 14.82 $/ha in Kentucky. Halich (2016) also stated that custom applicators in Kentucky charged 4.94 $/ha more for VRA (instead of URA) dry fertilizer. In this work it was assumed that the price that service providers charged is P^{VR}=22.39 $/ha for VRA and P^{UR}=$17.3 $/ha for URA. During TS, C_t^A is obtained by doing the area of the field times the P^{VR}, because VRA is needed to run the trial. During PS, when VRA was used, C_t^A is also calculated doing the area of the field times P^{VR}. In contrast, when URA was used, C_t^A is the area of the field times P^{UR}.

C_t^{CC} denotes the cost of soil sampling to obtain N$^{Apr.ss}$ and OM$^{Apr.ss}$. The cost of soil analysis in a laboratory in the Midwest was assumed to be $4/sample for N and $4/sample for OM$^{Apr.ss}$. In the high information sets, at the beginning of year 1, a soil sampling in the squared grid is performed both for N and OM, providing 32 samples with a cost of $8 each sample (since it was analyzed for both OM and N). Then, during TS a soil sample of N by plot was considered every year, composed of 110 samples at $4 each. The same criteria of one sample per plot for N was considered for the first year of PS, adding one sample for the border (total 111 samples). The soil sampling of OM was performed again in year 5 and 9 (every 4 years) always by the squared grid, providing 32 samples with a cost of $4 each sample. OM soil sampling cost was

not distributed over the 4 years, if not that it was computed at the moment it happened. For example, if L = 3:

- C_1^{SS} = 32 x 4 + 32 x 4 = \$256,
- $C_2^{SS} = C_3^{SS} = C_4^{SS}$ = 110 x 4 = \$440,
- C_5^{SS} = 32 x 4 + 32 x 4 = \$256,
- $C_6^{SS} = C_7^{SS} = C_8^{SS}$ = 32 x 4 = \$128,
- C_9^{SS} = 32 x 4 + 32 x 4 = \$256
- C_{10}^{SS} = 32 x 4 = \$128

The cost of N (C^N) was calculated by adding for the whole field the N used in each cell and multiplying by P^N. In the same way, revenue (**R**) is obtained by adding for the whole field the amount of grain harvested in each cell and multiplying by P^C.

$$C_t^N = P^N \sum_{j=1}^{256}(N_t \; x \; area_j) \tag{10}$$

$$R_t = P^C \sum_{j=1}^{256}(Y_t \; x \; area_j) \tag{11}$$

Finally, two summarizing variables were calculated. One is the profits (**Π**), calculated every year by subtracting the revenue and the different cost variables:

$$\Pi_t = R_t - C_t^N - C_t^F - C_t^A - C_t^{SS} \tag{12}$$

The other one is called present value (**PV**). It was calculated in April of year 1 (first day of the set) adding the discount value of all flows. For this, all costs were assumed to be incurred in April of each year, and revenue was obtained in October of each year (0.5 years after the income). Nominal values were discounted with an annual discount rate (r) of 5%.

$$PV_t = \frac{1}{(1+r)^{(t-1)}}[R_t] - \sum_{t=1}^{10} \frac{1}{(1+r)^{(t-0.5)}}[C_t^N + C_t^F + C_t^A + C_t^{SS}] \tag{13}$$

$$Annual \; PV = \frac{\sum_{t=1}^{10}(PV_t)}{10} \tag{14}$$

4.6 VALUE OF INFORMATION AND TECHNOLOGY

Throughout this work, it was assumed that all other management practices that are not related with N management are held constant. Also, it was assumed that, when no experiments were conducted over the field, the farmer used MRTN. Since MRTN does not take into account

28

site-specific information, the management plan in every cell will be the same, and the farmer will use URA. Following the methodology presented in section 3.2, the value of information and technology was calculated. Since the time when the farmer decides to change from running trials to regular production will affect the PV, for this calculation, it was assumed that the farmer would change at the optimal stopping time ($L_{z,I,T}^{OPT}$) for each z, I and T.

If trials were run over the field and data were used to select the best UR for the field, then the value of the information of those OFT ($V^{I.OFT}$) can be obtained by comparing the PV of the sets with $S^{low,URA}$ with the PV of the sets with $S^{no,URA}$ for each z:

$$V^{I.OFT}(z) = PV(z, L_{z,low,URA}^{OPT}, S^{low,URA}) - PV(z, 0, S^{no,URA}) \tag{15}$$

If trials were run and the information was used for VRA during the PS, then the realized value can be decomposed into two parts: the already explained $V^{I.OFT}$ and the added value of using VRT in the PS (after running trials) in the low information scenario ($V^{T.LOW}$). Based on the assumptions of this work, the farmer has no incentive to use VRA if no trials were run in the field, thus this value is only calculated as a complement of OFT. Then $V^{T.LOW}$ can be obtained by comparing the PV of the sets with $S^{low,VRA}$ with the PV of the sets with $S^{low,URA}$ for each z:

$$V^{T.LOW}(z) == PV(z, L_{z,low,VRA}^{OPT}, S^{low,VRA}) - PV(z, L_{z,low,URA}^{OPT}, S^{low,URA}) \tag{16}$$

If trials were run over the field together with the gathering of soil sampling information, then the realized value can be decomposed into two parts: the $V^{I.OFT}$ and the added value of the information from soil sampling ($V^{I.SS}$). Again, this value is always complementary to the $V^{I.OFT}$ because, in the assumptions of this work, the farmer would use a uniform MRTN rate if no trials were run in the field, thus there is no incentive to do soil sampling if no trials are run. Then $V^{I.SS}$ can be obtained by comparing the PV of the sets with $S^{low,URA}$ with the PV of the sets with $S^{high,URA}$ for each z:

$$V^{I.SS}(z) = PV(z, L_{z,high,URA}^{OPT}, S^{high,URA}) - PV(z, L_{z,low,URA}^{OPT}, S^{low,URA}) \tag{17}$$

Finally, if the information of the trials and soil sampling is used together with VRA, then the realized value can be decomposed into three parts: the $V^{I.OFT}$, the $V^{I.SS}$ and the value of VRT in the high information scenario ($V^{T.HIGH}$). This last value can be obtained by comparing the PV of the sets with $S^{high,VRA}$ with the PV of the sets with $S^{high,URA}$ for each z:

29

$$V^{T.HIGH}(z) = = PV\left(z, L_{z,high,VRA}^{OPT}, S^{high,VRA}\right) - PV\left(z, L_{z,high,URA}^{OPT}, S^{high,URA}\right) \qquad (18)$$

4.7 ENVIROMENTAL IMPACT

The environmental impact of N management is mostly related to the movement of N outside the crop system, producing contamination of sub-superficial water. This problem arises when precipitation events, NO_3-N availability and crop N-uptake do not agree in time (Dinnes et al, 2002). To compare the environmental impact that the different strategies would have, the N-leaching estimated by APSIM was considered. For each set (z,L,I,T) this variable was averaged by year and called annual N-leaching (ANL) (kg/ha).

$$ANL = \sum_{t=1}^{10} Nleaching_t \qquad (19)$$

The same comparisons explained in 4.7 to obtain the Value of the different information sources and technology were followed. In this case, instead of calling it value, since it did not involve economic variables, it was called N-leaching impact (NLI). In the same way, instead of using PV, the variable used was ANL. This way the environmental impact of on farm trials (NLI[.OFT]), of using VRA in low information scenario (NLI[.LOW]), of doing soils sampling (NLI[.SS]) and of using VRA in a high information scenario (NLI[HIGH]) were obtained using the following equations:

$$NLI^{I.OFT}(z) = ANL(z, L_{z,low,URA}^{OPT}, S^{low,URA}) - ANL(z, 0, S^{no,URA}) \qquad (20)$$

$$NLI^{T.LOW}(z) = ANL\left(z, L_{z,low,VRA}^{OPT}, S^{low,VRA}\right) - ANL\left(z, L_{z,low,URA}^{OPT}, S^{low,URA}\right) \qquad (21)$$

$$NLI^{I.SS}(z) = ANL\left(z, L_{z,high,URA}^{OPT}, S^{high,URA}\right) - ANL(z, L_{z,low,URA}^{OPT}, S^{low,URA}) \qquad (22)$$

$$NLI^{T.HIGH}(z) = ANL\left(z, L_{z,high,VRA}^{OPT}, S^{high,VRA}\right) - ANL\left(z, L_{z,high,URA}^{OPT}, S^{high,URA}\right) \qquad (23)$$

CHAPTER 5: RESULTS

This work involved a considerable amount of data. To show the results in a way that all the considerations made during this work can be seen, the approach of showing in detail only two S and then increase the number of sets and information was followed.

First, the breakdown of the economic variables of sets is shown. Then, those economic variables are summarized in the PV, with only one value per set and 21 sets, corresponding to the same z, were compared. Then the economic value and environmental value of the different sources of information and the technology was calculated. Finally, a data exploration was performed to understand what are the underlying causes of the different performance of the strategies under different weather scenarios.

5.1 RESULTS EXPLORATION

5.1.1 WEATHER CHARACTERIZATION

One of the main advantages of crop simulation is that allows to compare results of different N management strategies (combinations of information and technology) over different weather scenarios, maintaining all other factors equal (i.e. initial conditions of the field, prices). Since in this work weather is one of the main sources of variability and it will allow to explain differences in the results, in figure 4 the precipitation of each of the years explored in this work is shown, detailed in three different weather variables: pp^A, pp^S, pp^J. The first variable is called *Annual precipitation* and is the summation of the precipitation over the year. The following two variables area *Season* and *July's precipitation* and they were explained in section 4.4.1.
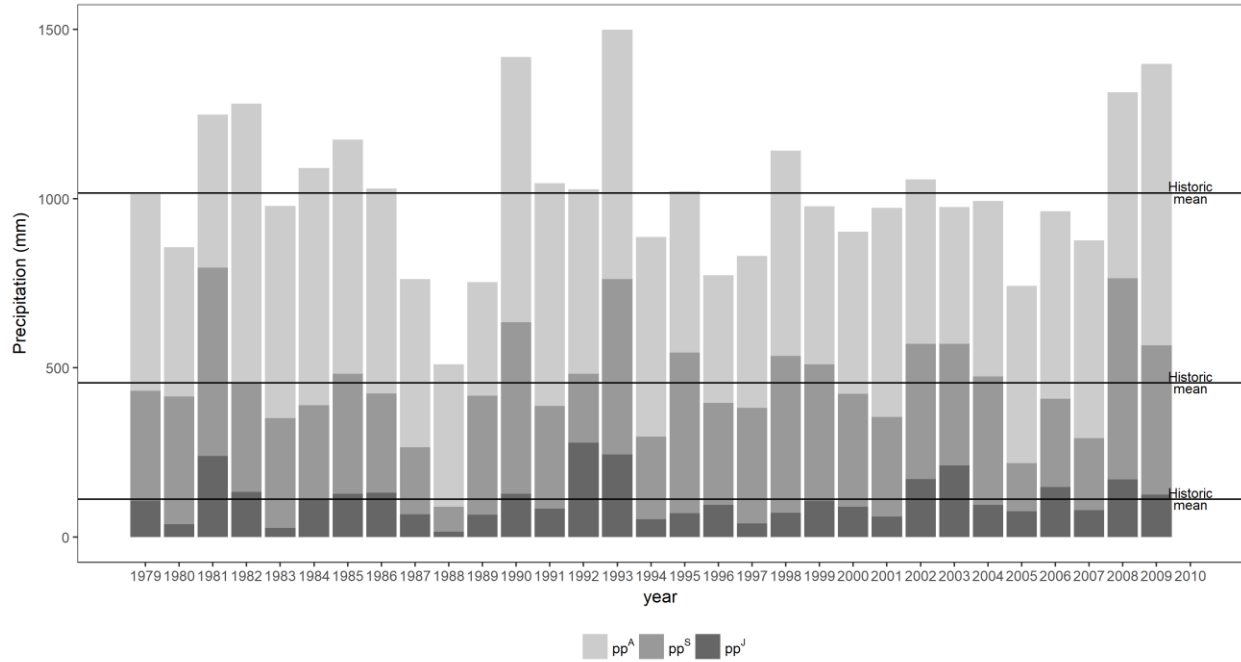
5.1.2 TWO SETS DESCRIPTION

To understand the economics considered in this work, the trajectory of revenues, profits and cost of two of the 10-year sets is showed in figure 3. From the 105 sets of this work, the chosen are set(1990,2,$S^{high,URA}$) together with Set(1990,0,$S^{no,URA}$). We will refer to them only by mentioning the S.

In figure 5 the total values for the whole field in ($/year) of the different variables are shown. In the first set, since L =2, there were two trials made before moving to regular production and then, in PS, URA was implemented. In $S^{no,URA}$, URA was used and no trials were involved. C^F was present only in the $S^{high,URA}$ and was higher for the first year and decreasing for the second since a learning curve was assumed for the farmer; after year 3 this cost is zero since no trials were run. As explained in section 4.3.1, the C^{SS} changed due to the change in the number of samples (depending if it is done by square grid or by plot) and the analysis performed (N every year, OM in year 1,5 and 9). C^A is the cost of application, it is higher in the $S^{high,URA}$ in year 1 and 2 since VRT was used, at P^{VR}. Then, both sets used URA at the lower P^{UR}.

The highest cost is the cost of fertilizer. This cost was constant for $S^{no,URA}$. In the $S^{high,URA}$ it is lower during the year 1 since trial rates are lower than MRTN. In year 2 it decreased even more, since the most profitable treatment was lower than MRTN, and in consequence the central

TR was lowered. When the field moved to the PS, the rates are estimated using the model and considering the $N^{Apr.ss}$, and they were lower than MRTN rate.

Moving to the top part of the figure, revenues were close between both sets during most of the years, being slightly higher for the $S^{no,URA}$ during years 1 to 4. Profits, on the other hand, started to be higher for the $S^{high,URA}$ after year 4, mainly due to the decrease in the use of fertilizer based on the better knowledge of the response curve.

Figure 5: Economic variables trajectory for two sets
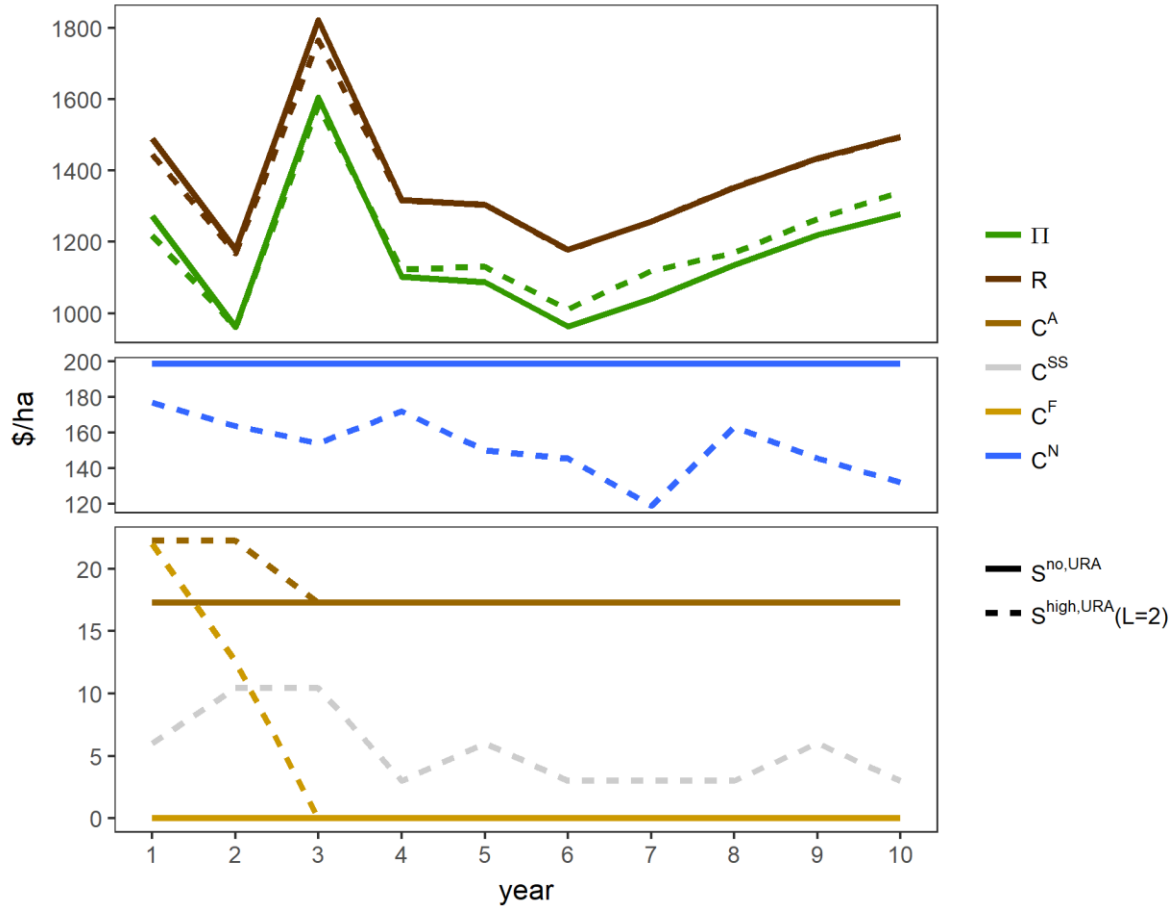
5.1.3 TWENTY-ONE SETS DESCRIPTION (z = 1990)

In Figure 6, all the sets for z=1990 are represented, showing the Annual PV of each strategy for increasing L. The four strategies that involve trials exceeded $S^{no,URA}$ for all the L options. Moreover, all the strategies had an optimal stopping time in $L^{OPT}=2$, where the Annual PV was maximized. The strategy with the highest PV was $S^{high,URA}$ with L = 2. This strategy overachieved the Annual PV of $S^{no,URA}$ by 16.79 \$/ha.
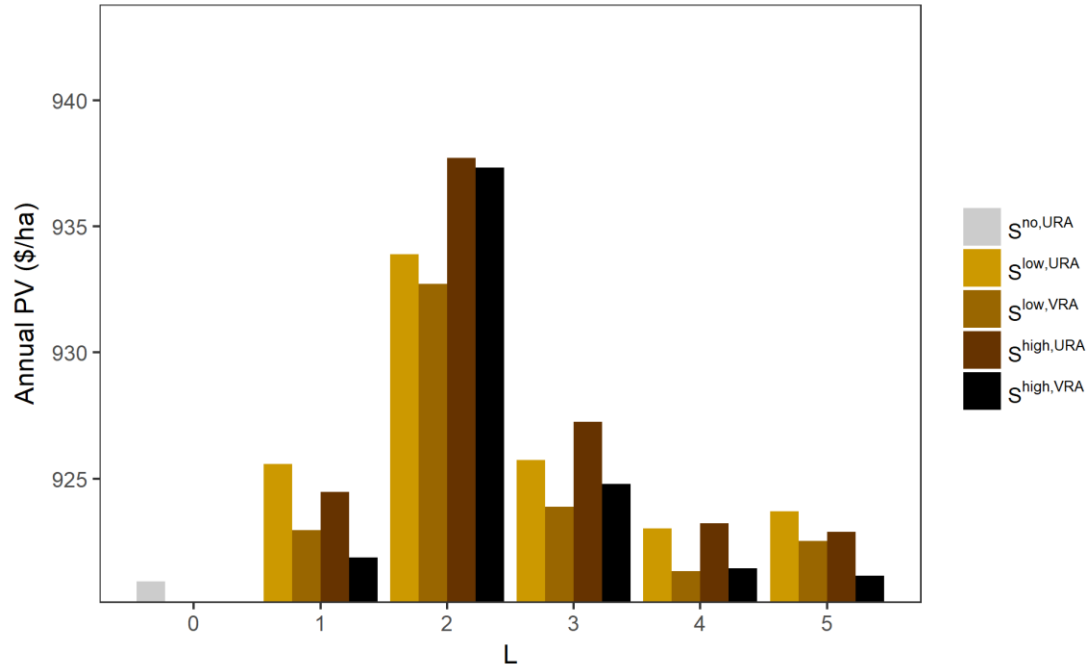
5.1.4 ALL SETS PV

Continuing with the pattern of increasing the number of sets shown, in Figure 7 the 105 simulated sets can be seen. They are grouped by z, and the **Annual ΔPV** is shown, which is the difference between the Annual PV of each of the strategies with the Annual PV of the strategy that does not involved trials for the same z.

This plot shows the variability found in the results of the simulation, and how the weather scenario could affect both the L^{OPT} and what strategy is the most profitable. Regarding the L^{OPT} different patterns can be seen. In z = 1980, the L^{OPT} change based on the information, being 1 year for the low information strategies ($S^{low,URA}$ and $S^{low,VRA}$) and 2 years for the high information ($S^{high,URA}$ and $S^{high,VRA}$). In z = 1985 the L^{OPT} was 1 year for all the strategies. In 1990 and 1995 the L^{OPT} was 2 for all the strategies. Simulations with z = 2000 showed a completely different pattern, where the new concept almost never overachieved the Annual PV of the $S^{no,URA}$, except for $S^{high,URA}$ L = 2 and L=4 with an Annual ΔPV of 0.89 \$/ha and for 0.43 \$/ha respectively.

On the other hand, the weather also affected the ranking of the strategies. Comparing all at the L^{OPT}, the best strategy was $S^{high,URA}$ for 4 of the 5 years, being 1980 the exception where $S^{low,URA}$ was the highest. Another noticeable pattern is how in z = 1985 and 2000 the high

34

information strategies overachieved considerably the low information strategies, while, in the other z, the difference is smaller.

Overall, as general insight from this graph, it can be said that in this field, the concept of OFT was more profitable that the concept of using MRTN in 4 of 5 weather scenarios, the optimal L was always between 1 and 2 and that $S^{high,URA}$ had the highest Annual PV in 4 of the 5 weather scenarios.

Figure 7: ΔPV for each strategy with increasing L by z. Annual ΔPV is the different of the Annual PV of each strategy with the Annual PV of $S^{no,URA}$ for the same z.
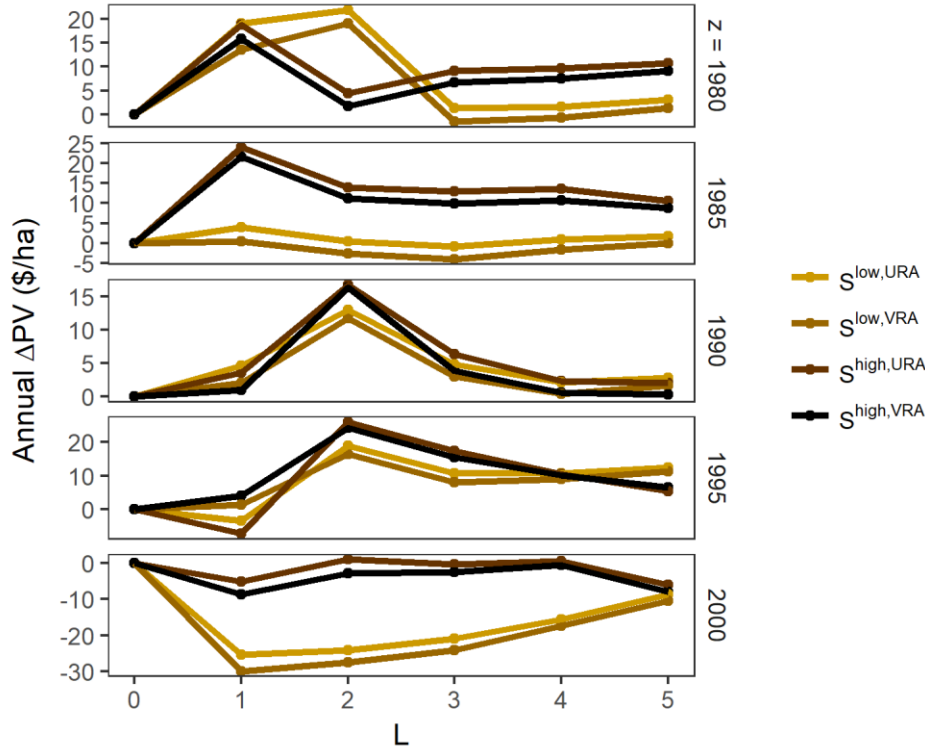
5.2 VALUE OF INFORMATION AND TECHNOLOGY AT L^{OPT}

The value of the four strategies is presented in table 2. These values are the annual ΔPV ($ / ha) using the L^{OPT} in each strategy and subtracting the corresponding annual PV ($/ha) of the $S^{no,URA}$ for the same z. In the last row, the values were averaged over the five weather scenarios.

Following the methods explained in section 4.7, $V^{I.OFT}$ is the value of the information obtained doing trials and it was 9.8 $/ha. $V^{T.LOW}$ was the added value of the VRT in a low information scenario (only with trial information and "free" variables) and it was -2.4 $/ha, meaning that using this technology after running trials did not pay the increased cost that the farmer faced to hire the variable rate equipment.

35

$V^{I.SS}$ is the added value of information from soil sampling and it was 7.4 \$/ha. That means that this source of information paid its cost and provided value for the farmer when PS strategy was URA. $V^{T.HIGH}$ was the added value of using VRT in the PS in a high information scenario and it was -1.8 \$/ha. That means that using VRT in post-production did not payed the cost with high information.

It is important to notice that even though the value of the technology was negative in both information level, this is the value in the second stage (post-trial production). VRT was valuably used in TS to run the trials and learn the production function. The 17.3 \$/ha of $S^{high,low}$ would not be possible without VR. Unfortunately, with the data we have it is not possible to isolate the value of VRT in TS, since it is also combined with all the factors involved in the proposed new concept of on-farm trial experimentation (trial design, analysis of the data, knowledge). Overall, in this work, VRT was valuable to run OFT, but not for doing a regular production using the information of those OFT. In this situation, the farmer should use/hire a VRT machine during the 1 or 2 years to run trials, and then use/hire a URA machine, at a lower cost.

Table 2: Left: Annual ΔPV (\$/ha). Right: the value of different sources of Information and technology.

z	$S^{low,URA}$	$S^{low,VRA}$	$S^{high,URA}$	$S^{high,VRA}$	$V^{I.OFT}$	$V^{T.LOW}$	$V^{I.SS}$	$V^{T.HIGH}$
1980	21.9	19.1	18.8	15.9	21.9	-2.8	-3.1	-2.9
1985	4.0	0.5	23.9	21.6	4	-3.5	19.9	-2.3
1990	13.0	11.8	16.8	16.4	13	-1.2	3.8	-0.4
1995	18.9	16.5	26.0	24.4	18.9	-2.4	7	-1.6
2000	-8.7	-10.6	0.9	-0.6	-8.7	-1.9	9.6	-1.5
AVERAGE	9.8	7.5	17.3	15.5	9.8	-2.4	7.4	-1.8

5.3 N-LEACHING TRAJECTORY AND ENVIRONMENTAL VALUE AT OPTIMAL STOPPING TIME

In Figure 8, the trajectory of the Annual N-Leaching (ANL) concerning L is shown. In most of the cases L = 5 is the number of trials that maximized the reduction in N-leaching. Nevertheless, in most of the weather scenarios, the big drop is produced at L^{OPT}, and the marginal contribution of higher L is insignificant. This allows explaining that, in this section, the NLC the different strategies will be calculated at the L^{OPT}. Note that L^{OPT} is an economic maximization, where the PV for each strategy was the highest. Moreover, since the graph shows that the contribution of L higher than L^{OPT} is small, the ANL at the economic L^{OPT} would be similar that the ANL at an environmental L^{OPT}.

Figure 8: Average Annual N-leaching trajectory at different L and weather scenarios
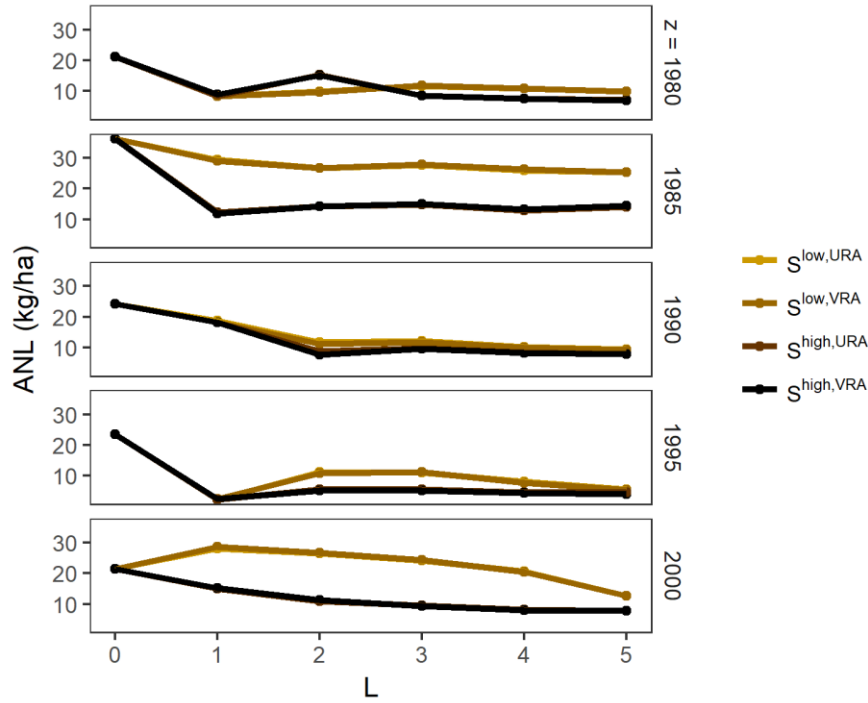
Contributing to show the environmental impact of OFE, Figure 9 shows the ANL for each strategy at the L^{OPT}. This figure is useful to perceive in total amounts the environmental benefit of the proposed concept. In most of the situations, the four strategies involving OFE reduced the ANL to half of the value of the $S^{no,URA}$.

Table 3 breaks down the total N-leaching amounts into specific contributions. On the left, the ΔNLC (reduction in N-leaching compared with the strategy of using MRTN) is shown. Then, it is decomposed into $NLC^{I.OFT}$, $NLC^{T.LOW}$, $NLC^{I.SS}$ and $NLC^{T.HIGH}$. The lowest the NLC, the lowest the environmental impact the strategy is producing. Doing OFT for L^{OPT} years and then using that information to optimize the UR in post-production provided a marginal $NLC^{I.OFT}$ of -10.4 kg/ha. Also, if the OFT information is used with VR, the marginal $NLC^{T.LOW}$ was -0.3 kg/ha, meaning that using VRT in post-trial production did not reduce the environmental impact considerably. $NLC^{I.SS}$ is the marginal impact of doing soil sampling with the OFT and it had a value of -5.9 kg/ha. $NLC^{T.HIGH}$ is the marginal environmental impact of using VRT in the post-trial production in a high information scenario and it was -0.8 kg/ha. That means that using VRT in post-production neither had a considerable reduction in N-leaching.

All in all, OFT together with soil sampling would reduce the annual N-leaching by 16.2 kg/ha. This is the $S^{high,URA}$, and it was also the strategy that maximized the PV in almost all the

weather scenarios (except one). Again, VRT is important in TS, for running trials that allows to reduce the N-leaching by knowing the yield response function. However, it did not contribute to reduce the N-leaching by itself in the post-trial production phase.
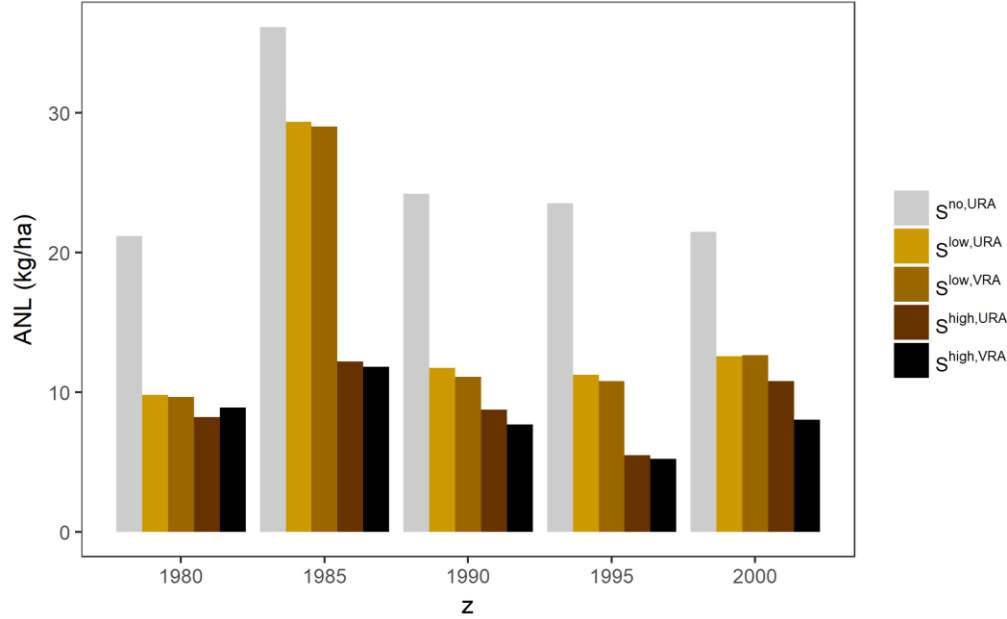
Figure 9: ANL of each strategy at L^{OPT}

Table 3: Left: Annual ΔANL (kg/ha). Right: NLI of Information and technology (kg/ha)

z	$S^{low,URA}$	$S^{low,VRA}$	$S^{high,URA}$	$S^{high,VRA}$	$NLI^{I.OFT}$	$NLI^{T.LOW}$	$NLI^{I.SS}$	$NLI^{T.HIGH}$
1980	-11.4	-11.5	-13.0	-12.3	-11.4	-0.1	-1.6	0.7
1985	-6.8	-7.1	-23.9	-24.3	-6.8	-0.3	-17.2	-0.4
1990	-12.5	-13.1	-15.5	-16.5	-12.5	-0.6	-3	-1
1995	-12.3	-12.7	-18.0	-18.3	-12.3	-0.4	-5.7	-0.3
2000	-8.9	-8.8	-10.7	-13.4	-8.9	0.1	-1.8	-2.8
AVERAGE	-10.4	-10.7	-16.2	-17.0	-10.4	-0.3	-5.9	-0.8

5.4 UNDERSTANDING UNDERLYING CAUSES: CASE ANALYSIS

One goal of this work is to detect general patterns that allow researchers to improve the new concept. Also, given the five different weather scenarios explored, another goal is to understand differences in those patterns, discuss the underlying causes and get insights that could help to adapt the new proposed concept to maximize profits in those situations.

To understand with more detail how different scenarios affected the profitability of the proposed new concept, especially patterns and exceptions observed in figure 7 were analyzed in

detail. The approach was to isolate those situations in cases and explore it using a useful graph with the important variables that explain the final results.

5.4.1 IMPORTANCE OF INFORMATION IN DIFFERENT WEATHER SCENARIOS

5.4.1.1 CASE 1: LOW PV OF OFE IN 2000

In figure 7 it was noticeable that in $z = 2000$, the proposed new concept did not increase the PV compared with the $S^{no,URA}$. To understand this exception, figure 10 shows the trajectory of different variables during the 10 years of two sets: set(2000, 0, $S^{no,URA}$) and set(2000, 1, $S^{high,URA}$). The top graph shows weather information, the middle one show economic variables and the bottom one show the **Cumulative ΔPV ($/ha)**. This variable shows the trajectory of the difference in PV between the most profitable set and the less profitable set of the graph.

It can be seen that there is a big difference in profits in year 1 (the trial year) where the yield of $S^{no,URA}$ overachieved the yield of the trial by 850 kg/ha, increasing revenues. Considering that the cost of the trial was also higher, the whole field profits were 145 $/ha higher. Afterwards, the higher efficiency of the $S^{high,URA}$ is noticed along the remaining years, showing higher profits and increasing the cumulative ΔPV especially in years 5 and 7. Nevertheless, this was not enough to compensate for the higher cost of the trial and the PV of the $S^{high,URA}$ was very close to the PV of $S^{no,URA}$.

Figure 10: Weather, economic variables and soil characteristics set(2000, 0, $S^{no,URA}$) and set(2000, 2, $S^{high,URA}$). Cumulative ΔPV is the cumulated difference in PV between the most profitable set and the less profitable set
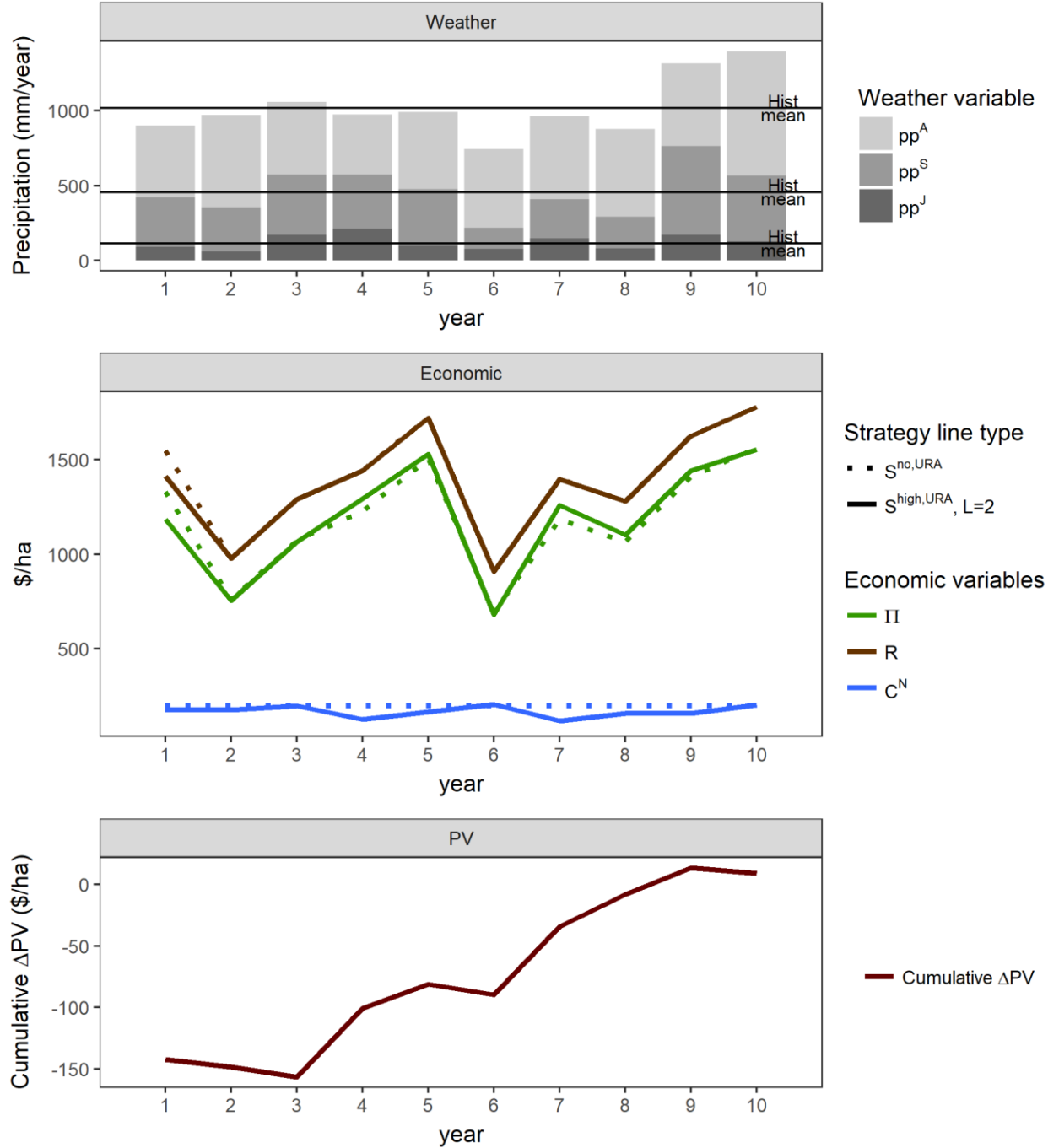
To illustrate the opportunity cost of the trial (OCT) (in $S^{high,URA}$) compared to $S^{no,URA}$, figure 11 is presented. For each weather scenario, for each trial year, the opportunity cost in $/ha was calculated subtracting the Profits of the set with trials with the Profits of the set without trial. The formula was:

$$OCT_{z,t} = \Pi_{z,t,high,URA} - \Pi_{z,t,no,URA}$$

It can be seen how, in fact, the opportunity cost of the first trial for z = 2000 was -150 $/ha. This cost is much higher than all other first trials and was not possible to be compensated with the afterward increased efficiency. To explain the underlying cause, Figure 12 shows the response of Y to N for the first trials in different *w*. Clearly, for z = 2000, the conditions of the field and the weather produced a high response to N. This means that the yield of the low rate treatments of the trial were low, causing that the average yield of the trial was lower than the one obtained with MRTN rate.

Possible solutions that should be discussed to avoid this problem could be test rates closer to MRTN, avoiding the low TR-90 kg/ha used in this work. Another one could be to have an unbalanced design, and instead of assigning a same number of plots to each treatment, assign less plots to the extreme treatments and more to the central ones. Another solution could be to make to make spread trials, with most of the field using the same rate and spread plots with treatments. This would reduce the area that is receiving extreme rates. All these solutions would improve the performance in a weather scenario like 2000, but could deteriorate the performance in the other scenarios. More research will be helpful to address what is the best overall strategy to assign treatments.

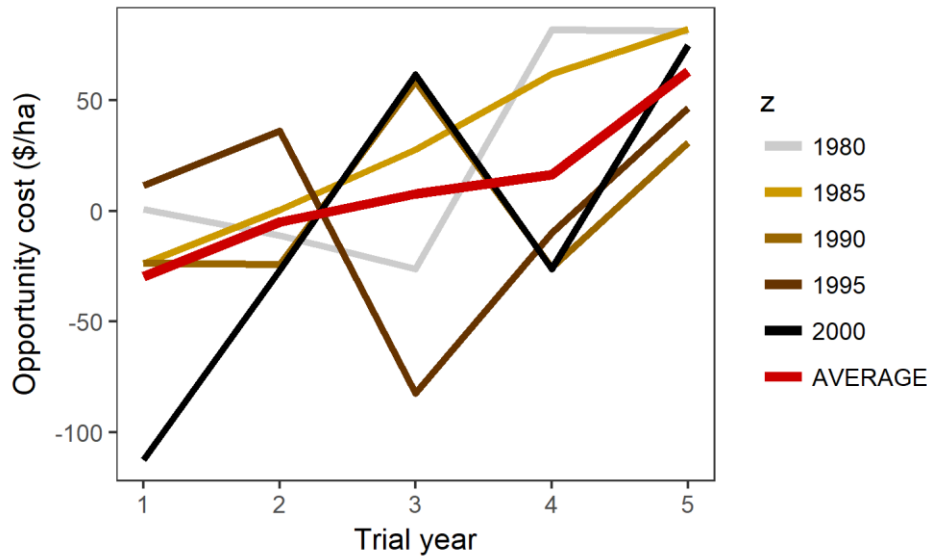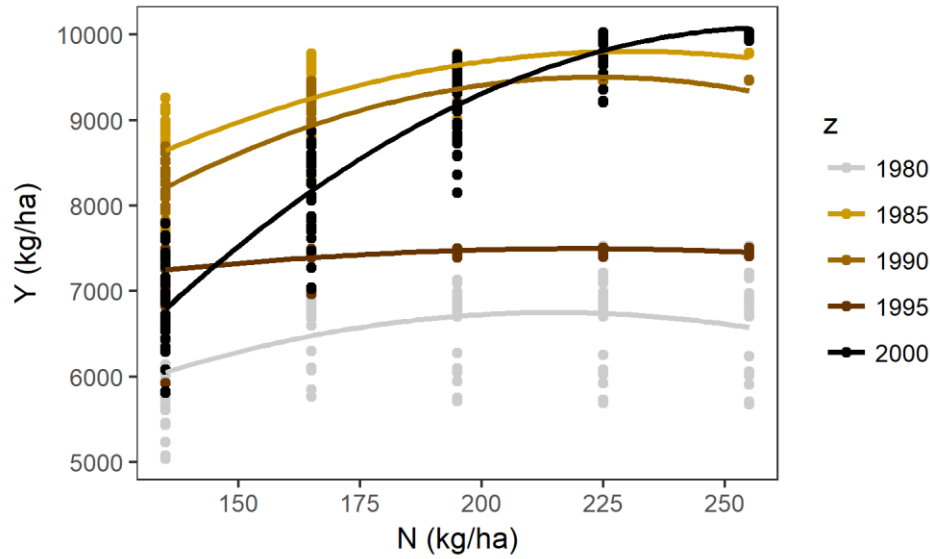*Figure 11: Annual opportunity cost of running trials compared with using MRTN rate*

41

Figure 12: Response of Y to N for the first trial on each Z

5.4.1.2 CASE 2: HIGH VS LOW INFORMATION IN z=1985 AND 2000

In figure 7, in the weather scenario of 1985 and 2000, there is a noticeable difference between the low information strategies ($S^{low,URA}$ and $S^{low,VRA}$) versus the high information strategies ($S^{high,URA}$ and $S^{high,VRA}$). To understand the underlying causes, a similar graph to the one showed in the previous analysis is shown in figure 13a. Two sets are included: set(1985,1, $S^{low,URA}$) and set(1985,1, $S^{high,URA}$), in representation of the low and the high information strategies. In the top graph it can be noticed that revenues are almost the same for both strategies. The difference is produced in the profits after year 4, mainly due to a lower fertilizer cost in $S^{high,URA}$. This behavior can be explained by looking at the weather and the N^{Apr} trajectory (lower graph). Year 4 was a dry year, thus the crop did not use the applied N, and this N stayed in the soil, increasing the N^{Apr} of year 5. $S^{high,URA}$ could capture this with soil sampling providing a lower N recommendation, reducing the cost and increasing profits. On the other hand, $S^{low,URA}$ did not include soils sampling information in the model. In consequence it was not able to capture the availability of N in the soil and the recommendations for year 5 were higher, increasing the cost and reducing profits. This produce that N^{Apr} was also high for year 6 and tend to stabilize to similar values in year 7.

In the same way, figure 13b shows the comparison between set(2000,1, $S^{low,URA}$) and set(2000,1, $S^{high,URA}$). It can be seen, again, that the increase in profits of the high information

strategy was due to the capacity of the model to detect high N^{Apr} in the soil and lower the N recommendations, increasing profits.

One of the objectives of this work was to test if past variables like Y/N_{t-1} could proxy the N in the soil. This case shows that the range of variation in Y/N_{t-1} explored during the trials did not provide enough information to build a model that could capture the excess of N after a dry year like 1988. On the other hand, in the high information model, the excess of N was well captured by soil sampling by adjusting the recommended N to the situation. Potentially, higher L would allow Y/N_{t-1} in the low information model to represent better the N^{Apr}, but as seen when comparing the PV, the L^{OPT} was always between 1 and 2, suggesting that it is economically more efficient to do soil sampling and decrease the number of trials.

Figure 13: Weather, economic variables and soil characteristics set(1985, 0, $S^{no,URA}$) and set(2000, 1, $S^{high,URA}$)
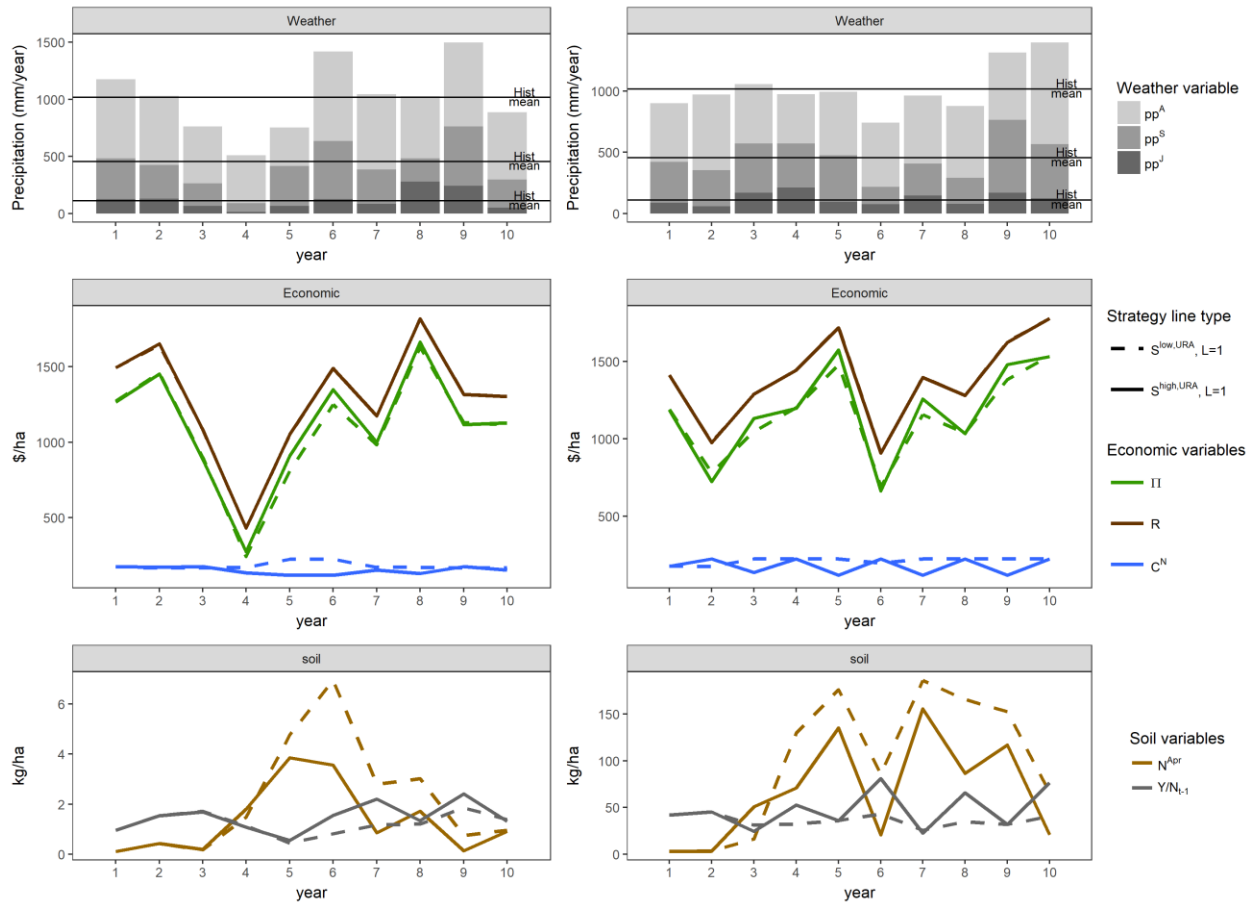
5.4.2 L^{OPT} CHANGE IN DIFFERENT WEATHER SCENARIONS

As explained in section 5.1, the L^{OPT} was in most of the situations between 1 and 2 years. Figure 7 shows how important would be to recognize when the number of trials is enough to provide profitable N management recommendations. In some situations, like in the weather scenario of 1980, increasing L one more year, from 1 to 2, reduced the Annual ΔPV of the high information Strategies to almost zero.

This section was attempted to discuss possible signs that could be considered by researchers and/or farmers to know that the trial data is enough and has sufficient quality and move to the PS. For this, the chosen strategy to analyze was $S^{high,URA}$, because it was the most profitable in all z, except one. Also, for simplification, the weather scenario z = 2000 was not considered in this analysis because, as shown in case 1, the concept of OFE was slightly profitable in that scenario and it could provide difficulties to this analysis.

## 5.4.2.1 CASE 3: $L^{OPT} = 1$ IN Z = 1980 (EXPERIENCING A WET WEATHER IN THE SECOND TRIAL)

In figure 7, in the weather scenario of 1980, there was a noticeable drop in the PV from L=1 to L=2. Figure 14a shows in detail more variables that help to explain the underlying causes of that drop. It can be seen that profits between both L in year 1 and 2 are similar, meaning that the second trial did not have a high opportunity cost compared to the regular crop using information of trial 1. It can also be seen that year 1 was below the historic average for the three weather variables (dry year) and year 2 was above average for the three weather variables (excessively wet year). In the model calculation using the weather weighting method (section 4.4) the second year showed a very uncommon weather and was weighted only 0.3%, being year 1 weighted 99.7%. Following the trajectory of the profits, L=1 starts to perform better after year 3, providing better recommendations that allowed to overcome L=2. That higher profitability in those years increased the cumulative

Since the model for L=2 weighted the second year fixed effects only by 0.3%, the recommendations should be very similar than the one with model L=1. Even more, the trial did not have a noticeable opportunity cost. Thus the model had more data at no cost. In that situation, why does L=1 provide better recommendations? Analyzing the models in detail, a possible explanation is that the fixed effects are only included as an interaction with N^T and N^{T^2}, but not

with the other variables (E, OM, Y/N$_{t-1}$). This was done to increase the range explored in those variables and have a model capable of performing in a wider range of situations. Nevertheless, in this case, the excessive wet year 2 is influencing the predictions through those other variables, increasing the recommended N. That means that even though the second trial had a low weight in the direct response to N, it had an indirect response through the coefficients that affect the other variables that interacted with N (for example through the term $\beta_{EN} * E_i * N_{t,i}^T$).

In summary, doing one more trial and getting information from a very excessive wet year harmed the N efficiency of the model. Considerations were made to avoid this using the weather weighting method of the year fixed effects, but it was produced indirectly. A possible solution would be to include fixed effects interactions with the other variables.

Figure 14: Weather, economic variables and soil characteristics. Cumulative ΔPV is the cumulated difference in PV between the most profitable set (L=1) and the less profitable set (L=2).
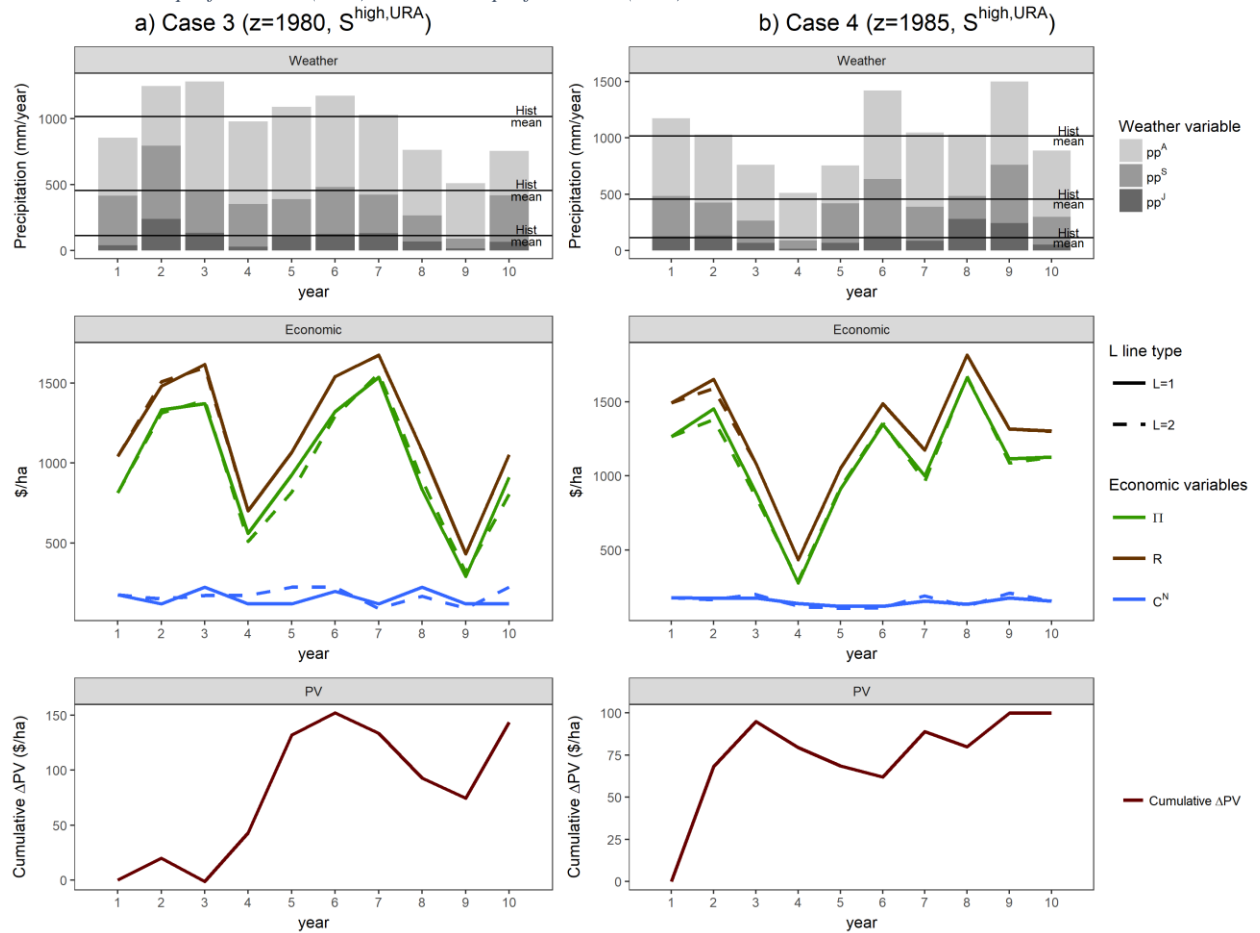
5.4.2.2 CASE 4: L^{OPT} = 1 IN 1985 (EXPERIENCING A SIMILAR WEATHER IN THE SECOND TRIAL)

In a similar way than in 1980, in the z = 1985, it is also a drop in the PV from L=1 to L=2. Figure 14b shows that the underlying causes were different in this scenario. Looking at the Cumulative ΔPV, the second trial in had an opportunity cost compared with regular crop using information of only one trial. The final Cumulative ΔPV between L=1 and L=2 was 100 $/ha and most of that difference (68.25 $/ha) was built in this second year. Moreover, this high opportunity cost, the second trial had similar weather in the growing season than the first, and thus did not provide new information to the model. Following the time line, in year 3 again the L=1 overperformed the L=2 model, by providing a slightly lower N recommendation that decreased the cost of fertilizer and produced higher profits. These three first years have a higher impact on the PV since the discount rate is lower and most part of the final PV difference is built. Afterwards, the model switched on their performance over the years, slightly loosing and then recovering the Cumulative ΔPV.

5.4.2.3 CASE 5: L^{OPT} = 2 IN 1990 (EXPERIENCING A WET SEASON IN THE FIRST TRIAL)

In the z = 1990, the L^{OPT} was 2. Figure 15a shows the underlying causes of this difference. In this case, the first year happened to be excessively wet. This had two consequences. First, the N recommendations for L=1 were high, based on the expected high response obtained in a trial during a wet year. Second, the opportunity cost of doing one more trial was low, since the recommended rates with L=1 were high for the weather experienced in the year, the rates of the trial were more efficient. Year 3 was slightly wet, especially in July and thus the model with L =1 performed better, recovering the difference of the second year. Afterwards, the L=2 was more efficient, building slowly a final Cumulative ΔPV of 132 $/ha

Figure 15: Weather, economic variables and soil characteristics. Cumulative ΔPV is the cumulated difference in PV between the most profitable set (L=2) and the less profitable set (L=1).
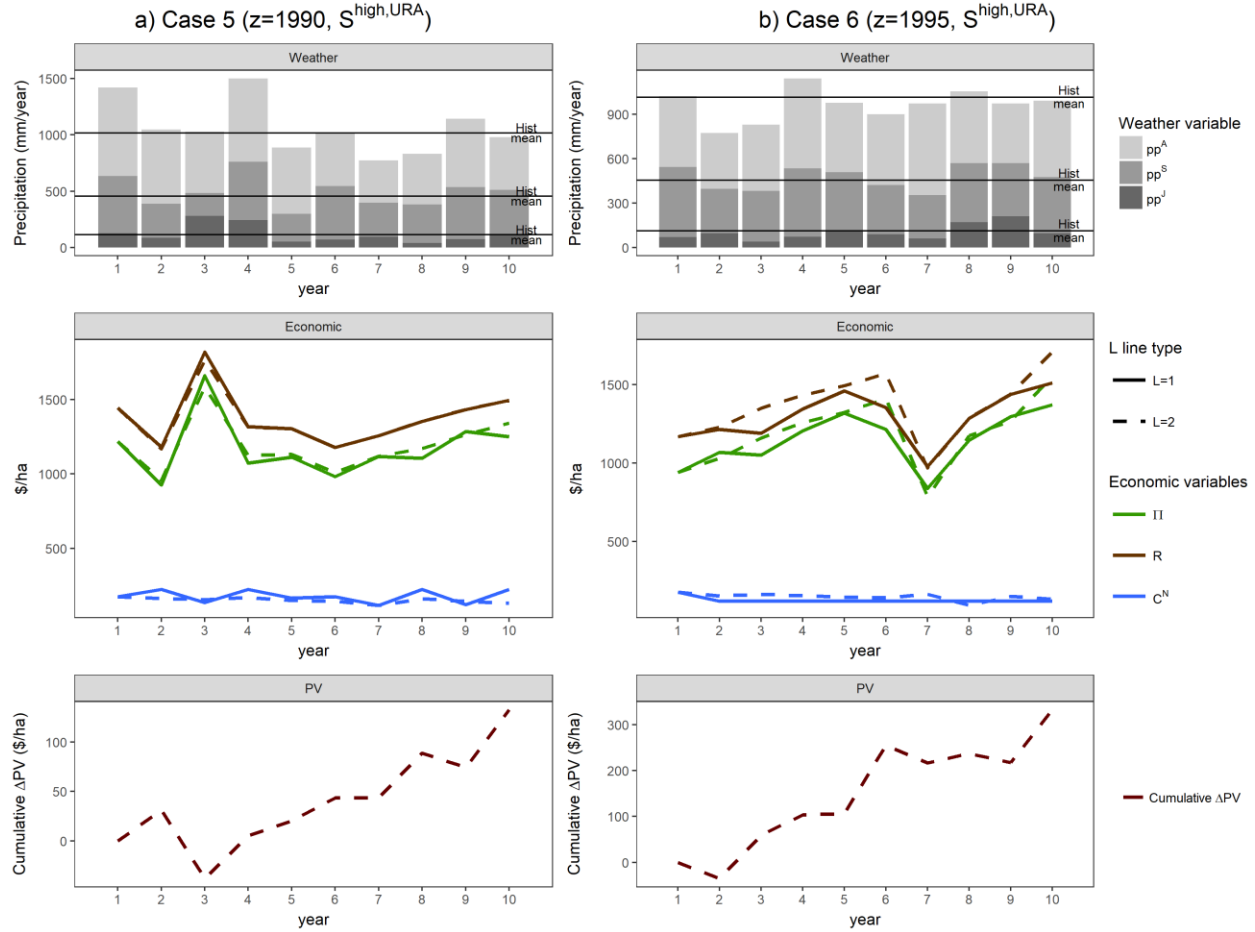
5.4.2.4 CASE 6: $L^{OPT} = 2$ IN 1995 (EXPERIENCING A DRY CRITIC PERIOD IN THE FIRST TRIAL)

In the $z = 1995$, the L^{OPT} was 2. To understand the reasons, figure 15b shows the graph of meaningful variables. July's precipitation in the first trial was low. July is the time when the crop is going through the critical period, and water limitations in this stage caused that the response of yield to N in trial one was low (figure 16). The second trial experienced a weather closer to the average weather, regarding July's precipitation and Season's precipitation. Also, the opportunity cost of this trial compared with L=1 strategy was low. Since trial one had a low response to N, the recommended N rates were low. Trial two allowed to explore a weather closer to the average with a low cost of opportunity, and the recommended N rates were higher and more profitable.

47

Afterwards, the L=2 strategy was more efficient, accumulating slowly after year 3 a final PV difference of 332.09 $/ha.

Figure 16: Response of Y to N for the first and the second trials in z = 1995.
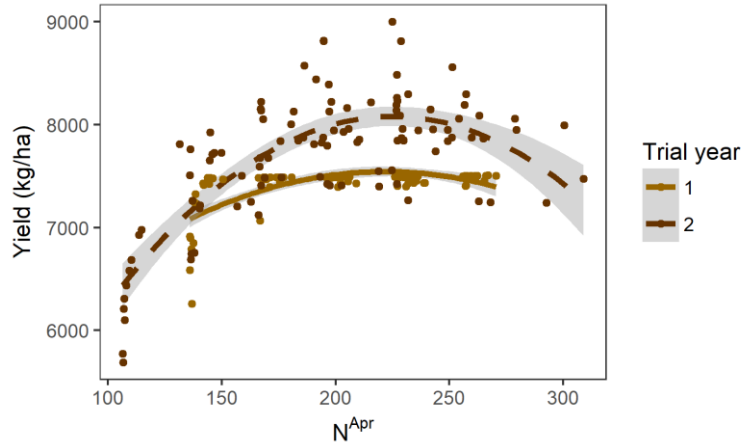
5.5 INSIGHTS FOR FINDING THE OPTIMAL STOPPING TIME EX-ANTE

Some general patterns discussed in previous sections included that L was most of the times between 1 and 2. In the section 5.4.2 in cases 3 to 6, differences in L^{OPT} were analyzed in detail, to understand the underlying causes. The weather scenarios 1980 and 1985 had $L^{OPT} =1$. The one from 1980 because the second trial experienced high precipitations and in consequence predicting high N rates (due to an indirect effect in the model). The other was caused by a second trial experiencing a similar weather than the first trial, having an opportunity cost and not providing more information.

The weather scenarios of 1990 and 1995 had $L^{OPT}= 2$. The one from 1990 was produced by a wet season in the first trial, producing high N recommendations that were fixed by including a second trial. The other one was produced by a dry critical period, producing low N recommendations that were fixed again by including a second trial.

The ΔPV presented in Figure 7 allowed to detect how important would be for the farmer and researcher to have insights to detect the optimal stopping time. The PV between L=1 and L=2 can change noticeable, and sometimes running one more year of experiments could damage a big part of the benefit of the proposed new concept of OFE. The goal of this section is to compare different z to detect if there are rules that could help the researcher to recognize that the available data is enough to provide profitable N management recommendations.

The variables shown in Table 4 were selected with the aim of recognizing the L^{OPT}. The main conclusions are:

-$Y^{max.r}$: is the highest Y of the trial compared with the average historic mean Y for the period 1980-2009 using MRTN. The attempt is to characterize the potential yield of the trial -without N limitations- in relation with the historical average yield. The hypothesis was that high or low yields in the first trial will increase the importance of the second trial. Based on the table, this happened in z=1995 were the first trial was 0.86 and the second 1.07. On the other hand, it did not happen in z=1980 were the first trial had a similar $Y^{max.r}$ than in 1995 but the L^{OPT} was not shifted to the second trial.

-N^{Apr}: is the average N^{Apr} in the field. The hypothesis was that if N^{Apr} was high, the response of Y to N in the trial would be low, and then the recommended rates will be low. A second trial would be valuable in this situation. Considering that the EONR (using N^T) ranged between 160-260 kg/ha (depending on the year), and that the lowest treatment was 90 kg below the TR (224 kg/ha the first year), the N^{Apr} range showed in the table was not high enough to avoid the trial explore rates close to the EONR.

-CP: pp^A, pp^S and pp^J precipitations are shown in the table. Also, following the equations in section 4.4.1, pp^S and pp^J are transformed probability (following a normal distribution) and then the combined probability was calculated (**CP**) using the following formula:

$$CP_t = PZ_t^S \, PZ_t^J \qquad\qquad\qquad (24)$$

This variable shows how close to the historic weather average are the season and the July precipitation together. A very unusual value of any of them will make the CP low. The table shows promising expectations for this value to predict the L^{OPT}. In the weather scenario of 1985 a trial with CP = 0.15 maximized profits. In weather scenarios of 1990 and 1995 the initial CP was low (0.04 and 0.08 respectively) and a second trial maximized profits. The weather scenario of 1980 was unusual because the CP of the first trial was low, and the second was even lower. In this case the first trial provided higher profits. More exploration needs to be done in z = 2000. The first trial had the higher CP. Nevertheless, the response of Y to N was much higher than in other trials, producing high N recommendations. This was fixed by including one more trial. This leads to think that some more weather variables should be weighted besides pp^J and pp^S.

Y/N^{overlap}: This value is showing how much of a normal range of variation for Y/N$_{t-1}$ was captured by the model. Using historic data of corn in the field using MRTN the 10% and 90% percentile of the Y/N$_{t-1}$ was set in 26 and 45 respectively. Then, the percentiles of the accumulated Y/N$_{t-1}$ by the model were obtained (accumulated means that the second trial year also includes the range explored by the first year). Finally, it was calculated how much of the historical range was explored by the trial (for example, if the percentiles for the trial were 30 and 44, the overlap would be 14/(45-26)=73%. The hypothesis was that a low overlap in the first trial would make the second trial more valuable. Clearly, more trials allowed to explore higher overlaps, but not that a low overlap is reason sufficient to make another trial, since years like 1980 and 1985 had <0.12 in the first trial and that was also the L^{OPT}.

*Table 4: Variables describing the conditions for each trial. * shows the L^{OPT}*

z	year	Y^{max}	Y$^{max.r}$	N^{Apr}	pp^S	pp^J	PZ^S	PZ^J	CP	Y/N^{overlap}
1980	*1	7531	0.87	2.61	415	38	0.40	0.14	0.06	0.04
1980	2	9975	1.15	22.61	796	240	0.01	0.01	0.00	0.39
1985	*1	9786	1.13	3.67	482	127	0.43	0.36	0.15	0.12
1985	2	10691	1.23	13.35	424	131	0.42	0.34	0.14	0.76
1990	1	9471	1.09	17.23	635	127	0.12	0.36	0.04	0.13
1990	*2	8749	1.01	8.26	387	84	0.33	0.37	0.12	0.70
1995	1	7507	0.86	6.64	545	70	0.28	0.29	0.08	0.16
1995	*2	9273	1.07	32.18	396	95	0.35	0.44	0.15	0.34
2000	1	10032	1.16	2.89	423	89	0.42	0.40	0.17	0.04
2000	*2	7720	0.89	3.36	355	60	0.26	0.23	0.06	0.23

CHAPTER 6: SUMMARY AND CONCLUSIONS

This study analyzed the economic optimization together with the environmental trade-offs of different strategies for conducting on-farm experimentation and how the results changed with different weather scenarios. Simulations were conducted in order to answer the initial five questions.

The first question that the work targeted was if the concept of OFE was profitable. Simulation results showed that the concept was profitable, if the right strategy and optimal stopping time was selected. At the optimal stopping time, the value of running trials together with measuring the "free" variables was 9.8 $/ha. The additional value of gathering soil sampling information at the same time was 7.4 $/ha. VRA was necessary for running trials; nevertheless, using VRA in the after-trial stage was not cost-beneficial, having a negative value of -2.4 $/ha (with low information) and -1.8 $/ha (with high information).

The second question investigated was what variables should be measured over the field and incorporated into the model that would maximize profits. Results showed that the "free" variables (Elevation, Y/N_{t-1}, as-applied N, Y) together with soil sampling information of N^{Apr} and OM^{Apr} following a soil sampling strategy (every year for N^{Apr} and every four years for OM) covered the cost and maximized profits for the farmer. The benefit of soil sampling was noticeable when residual N in the soil was high, such as when there was a dry year with low Yield or when a high response of yield to N provided high N recommendations. The model with soil sampling could capture that condition and decrease the following recommended N rate, increasing profits. On the other hand, the model without soil sampling information recommended high N rates in those situations, decreasing profits and sometimes making them negative (like in z=1985 and 2000). Y/N_{t-1} attempted to capture this condition, but since the optimal number of trials was one or two, the model was built with low variation in this variable, and it was not able to capture extreme situations that led to an increase in residual N. Since trials have an opportunity cost, this was not economically efficient, and it was more profitable to practice soil sampling.

The third question examined was how many years to run trials (L^{OPT}) before moving to regular production using the results of those trials. Simulations showed that the L^{OPT} was 1 or 2 in all the scenarios. The underlying causes were explored. The reasons that made L^{OPT} to be 1

were that the second trial explored a more unlikely weather (in terms of distance to the historic average of pp^S and pp^J) than the first one, or a weather similar to the first one (consequently, not adding information). The reasons that made L^{OPT} to be 2 were that the first trial explored an unlikely weather, and thus adding one more trial improved the recommendations.

The fourth question was related to the environmental benefits of OFE. This question was studied by analyzing the possible reduction in N-leaching that could be produced by the different combinations of L, technology, and information. The first conclusion is that the L^{OPT} that maximized profits also produced a high drop in the N-leaching, and thus doing more trials did not have a significant impact in the N-leaching reduction. This means that being economically efficient and stopping trials at L^{OPT} also reduced N-leaching as much as possible. The second conclusion was related to the breakdown of the impact of the Information and Technology over the N-leaching reduction. Simulations showed that running OFT decreased N-leaching by 10.4 kg/ha on average. Doing soil sampling together with OFT decreased the N-leaching another 5.9 kg/ha. Using VRA decreased N-leaching by 0.3 kg/ha and by 0.8 kg/ha in the low information and in the high information scenarios respectively. This means that VRA in the post-trial stage did not have a noticeable effect in the reduction of N-leaching.

The fifth question was what insights can be used to detect the L^{OPT} in ex-ante situations. Simulations showed how important this decision would be, since often doing one more trial could reduce the whole benefit of the OFE. Variables that described conditions explored during the trial were used to detect possible insights that could lead to detect when a trial would provide profitable N management advice. Simulations showed promising results for a variable that assigns a combined probability to precipitations during the season and during July based on historic weather. Tentatively, a combined probability of 0.15 in the weather of the first trial suggested the optimal stopping time. If the probability is lower, another trial is needed. However, this rule needs to be explored in more weather scenarios.

One conclusion of the work that was not an initial question is related to the Treatment's selection strategy. In year 2000, the new concept of OFE did not increase profits compared with not running trials and using MRTN. The reason was a high opportunity cost of the trial. This was produced due to a combination of low N^{Apr} with a high response of yield to N. In that context, the low rates of the trial had a low Yield, decreasing the profitability. This result suggested that

the cost of the trial is important, and that more research should be done to improve the Dynamic Treatments assignment used in this work. Three possible options to decrease the opportunity cost of the trial were suggested. The first was to concentrate the treatments closer to the MRTN. Another option was to have unbalanced designs, where a lower number of plots are assigned to treatments with extreme values (low and/or high) and a higher number of plots are assigned to the central treatments. The last option was to, instead of running whole field trials, spread out plots in the field to avoid having large parts of the field with rates that could be lower than the optimal.

OFE is a complex economic problem that involves decisions about how to run the trials (whole field, part of the field), how to assign the treatments (what rates to explore in the trial), what variables should be measured ("free" variables, soil sampling variables), when to stop running trials (L^{OPT}), how to analyze the data (OLS, panel spatial error model), how to calculate predictions with more than one year of data (equally weighted, weather weighted). Moreover, results are affected by the weather of the year and by the field conditions, providing a source of random variability that makes the problem more complex. Given the complexity of the problem and the amount of possible combinations of decisions, not all of them were tested in this work. Some of the decisions were addressed with a research approach of testing different combinations (like what variables should be included and different L). Others were decided based on experience or previous works (doing whole field trials, the treatment selection strategy, using SER and obtaining predictions weighting the trial data based on their weather probability). This study could be extended by testing all those decisions that were taken as assumptions in this work. It could also be extended by testing more weather scenarios and fields.

The results of the simulations in this work provided helpful information that underscores the exciting possibilities that exist related to OFE. After the work, we truly consider that the complexity of the problem makes Crop Simulation a valuable tool to start thinking about all the possible ways to optimize OFE. Crop simulation allows researchers to create many equal fields and apply different strategies and weathers, testing combinations that would not be possible to test in real situations. In real situation, if one strategy is tested in a site, another one cannot be tested in the same site. It also allows to test these combinations at low cost, saving time and effort. The trends observed in simulation results are a sound foundation that should be

53

complemented with empirical data and used together to optimize the actual protocol of the DIFM project.

REFERENCES

Anselin, L., Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). A spatial econometric approach to the economics of site-specific nitrogen management in corn production. American Journal of Agricultural Economics, 86(3), 675-687.

Archontoulis, S.V., Miguez, F.E., and Moore, K.J.(2014a). Evaluating APSIM maize, soilwater, soilnitrogen, manure, and soil temperatures modules in the Midwestern United States. Agron.J. 106, 1025–1040.doi: 10.2134/agronj2013.0421

Bakhsh, A., Kanwar, R. S., Jaynes, D. B., Colvin, T. S., & Ahuja, L. R. (2001). Simulating effects of variable nitrogen application rates on corn yields and NO3–N losses in subsurface drain water. Transactions of the ASAE, 44(2), 269.

Berck, P., & Helfand, G. (1990). Reconciling the von Liebig and differentiable crop production functions. American Journal of Agricultural Economics, 72(4), 985-996.

Bivand, R., Hauke, J., & Kossowski, T. (2013). Computing the J acobian in G aussian Spatial Autoregressive Models: An Illustrated Comparison of Available Methods. Geographical Analysis, 45(2), 150-179.

Bivand, R., & Piras, G. (2015). Comparing implementations of estimation methods for spatial econometrics. American Statistical Association.URL http://www.jstatsoft.org/v63/i18/.

Brorsen, B. W., & Richter, F. G. C. (2012). Experimental designs for estimating plateau-type production functions and economically optimal input levels. Journal of Productivity Analysis, 38(1), 45-52.

Bullock, D. S. (2013, June). Simulating the value of information generated by on-farm agronomic experimentation using precision agriculture technology. In Proceedings of IATRC Symposium on Productivity and Its Impacts on Global Trade.

Bullock, D. G., & Bullock, D. S. (1994). Quadratic and quadratic-plus-plateau models for predicting optimal nitrogen rate of corn: A comparison. Agronomy Journal, 86(1), 191-195.

Bullock, D.G., D.S. Bullock, E.D. Nafziger, T.A. Doerge, S.R. Paszkiewicz, P.R. Carter,and T.A. Peterson. 1998. "Does variable rate seeding of corn pay?" Agronomy Journal 90:830–836.

55

Bullock, D. G., & Bullock, D. S. (1994). Quadratic and quadratic-plus-plateau models for predicting optimal nitrogen rate of corn: A comparison. Agronomy Journal, 86(1), 191-195.

Bullock, D. S., Lowenberg-DeBoer, J., & Swinton, S. M. (2002). Adding value to spatially managed inputs by understanding site-specific yield response. Agricultural Economics, 27(3), 233-245.

Bullock, D. S., & Lowenberg-DeBoer, J. (2007). Using spatial analysis to study the values of variable rate technology and information. Journal of Agricultural Economics, 58(3), 517-535.

Bullock, D. S., Ruffo, M. L., Bullock, D. G., & Bollero, G. A. (2009). The value of variable rate technology: an information-theoretic approach. American Journal of Agricultural Economics, 91(1), 209-223.

Casanoves, F., Macchiavelli, R., & Balzarini, M. (2007). Models for multi-environment yield trials with fixed and random block effects and homogeneous and heterogeneous residual variances. The Journal of Agriculture of the University of Puerto Rico, 91(3-4), 117-131.

Castleberry, R. M., Crum, C. W., & Krull, C. F. (1984). Genetic Yield Improvement of US Maize Cultivars under Varying Fertility and Climatic Environments 1. Crop science, 24(1), 33-36.

Chambers, R. G., & Lichtenberg, E. (1996). A nonparametric approach to the von Liebig-Paris technology. American Journal of Agricultural Economics, 78(2), 373-386.

Dinnes, D. L., Karlen, D. L., Jaynes, D. B., Kaspar, T. C., Hatfield, J. L., Colvin, T. S., & Cambardella, C. A. (2002). Nitrogen management strategies to reduce nitrate leaching in tile-drained Midwestern soils. Agronomy journal, 94(1), 153-171.

Frank, M. D., Beattie, B. R., & Embleton, M. E. (1990). A comparison of alternative crop response models. American Journal of Agricultural Economics, 72(3), 597-603.

Millo, G., & Piras, G. (2012). splm: Spatial panel data models in R. Journal of Statistical Software, 47(1), 1-38. URL http://www.jstatsoft.org/v47/i01/.

Grimm, S. S., Paris, Q., & Williams, W. A. (1987). A von Liebig model for water and nitrogen crop response. Western Journal of Agricultural Economics, 182-192.

Haber, F. 1905. Thermodynamics of technical gas reactions http://nobel.sdsc.edu/laureates/chemistry-1918-1-bio.html[Accessed 2006]

Halich, G. 2016. Custom Machinery Rates Applicable to Kentucky (2016). Lexington, Kentucky, March.

Heady, E. O., Jacobson, N. L., Freeman, A. E., & Madden, J. P. (1964). Milk production functions incorporating variables for cow characteristics and environment. Journal of Farm Economics, 46(1), 1-19.

Holzworth, D. P., Huth, N. I., Zurcher, E. J., Herrmann, N. I., McLean, G., Chenu, K., ... & Brown, H. (2014). APSIM–evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software, 62, 327-350.

Llewelyn, R. V., & Featherstone, A. M. (1997). A comparison of crop production functions using simulated data for irrigated corn in western Kansas. Agricultural Systems, 54(4), 521-538.

Lowenberg-DeBoer, J. (2003, February). Precision farming or convenience agriculture. In Solutions for a better environment–Proceedings of the 11th Australian Agronomy Conference'. Geelong, Vic.(Australian Society of Agronomy)

Ma, B. L., & Dwyer, L. M. (1998). Nitrogen uptake and use of two contrasting maize hybrids differing in leaf senescence. Plant and soil, 199(2), 283-291.

Miao, Y., Mulla, D. J., Batchelor, W. D., Paz, J. O., Robert, P. C., & Wiebers, M. (2006). Evaluating management zone optimal nitrogen rates with a crop growth model. Agronomy Journal, 98(3), 545-553.

Miller, A. 2013. 2013 Indiana Farm Custom Rates. West Lafayette, Indiana, June.

Olson, R.A., Dreier, A.F., Thompson, C., Frank, K. and Grabouski, P.H. 1964 Using fertilizer nitrogen effectively on grain crops. Nebr. Agric. Exp. Stn Bull. SB 479. 42 pp.

Pan, W.L., D.R. Huggins, G.L. Malzer, C.L. Douglas, Jr., and J.L. Smith. 1997. Field heterogeneity in soil–plant nitrogen relationships: Implementations for site-specific management. p. 88–99. In F.J. Pierce and E.J. Sadler (ed.) The state of site-specific management for agriculture. ASA, CSSA, and SSSA, Madison, WI.

57

Paris, Q. (1992). The von Liebig hypothesis. American Journal of Agricultural Economics, 74(4), 1019-1028.

Peralta, N. R., Costa, J. L., Balzarini, M., & Angelini, H. (2013). Delineation of management zones with measurements of soil apparent electrical conductivity in the southeastern pampas. Canadian Journal of Soil Science, 93(2), 205-218.

Plastina, A., A. Johanns, and J. Erwin. 2016. 2016 Iowa Farm Custom Rate Survey. Ames, Iowa, March.

Team, R. C. (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/.

Sawyer, J., Nafziger, E., Randall, G., Bundy, L., Rehm, G., & Joern, B. (2006). Concepts and rationale for regional nitrogen rate guidelines for corn.

Shrader, W. D., Fuller, W. A., & Cady, F. B. (1966). Estimation of a Common Nitrogen Response Function for Corn (Zea mays) in Different Crop Rotations 1. Agronomy Journal, 58(4), 397-401.

Hexem, R. W., Sposito, V. A., & Heady, E. O. (1976). Application of a two-variable Mitscherlich Function in the analysis of yield-water-fertilizer relationships for corn. Water Resources Research, 12(1), 6-10.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for Effingham County, Illinois. Available online. Accessed Nov-2017.

Stein, D. 2014. 2014 Custom Machine and Work Rate Estimates. East Lansing, Michigan, November.

Swanson, E. R., Taylor, C. R., & Welch, L. F. (1973). Economically optimal levels of nitrogen fertilizer for corn: An analysis based on experimental data, 1966-1971. Illinois Agricultural Economics, 16-25.

Swinton, S. M., & Lowenberg-DeBoer, J. (1998). Evaluating the profitability of site-specific farming. Journal of production agriculture, 11(4), 439-446.

Swinton, S. M., & Lowenberg-DeBoer, J. (1998). Evaluating the profitability of site-specific farming. Journal of production agriculture, 11(4), 439-446.

Tembo, G., Brorsen, B. W., Epplin, F. M., & Tostão, E. (2008). Crop input response functions with stochastic plateaus. American Journal of Agricultural Economics, 90(2), 424-434.

Thorp, K. R., Batchelor, W. D., Paz, J. O., Steward, B. L., & Caragea, P. C. (2006). Methodology to link production and environmental risks of precision nitrogen management strategies in corn. Agricultural Systems, 89(2-3), 272-298.

Thorp, K. R., DeJonge, K. C., Kaleita, A. L., Batchelor, W. D., & Paz, J. O. (2008). Methodology for the use of DSSAT models for precision agriculture decision support. computers and electronics in agriculture, 64(2), 276-285.

Tumusiime, E., Wade Brorsen, B., Mosali, J., Johnson, J., Locke, J., & Biermacher, J. T. (2011). Determining optimal levels of nitrogen fertilizer using random parameter models. Journal of Agricultural and Applied Economics, 43(4), 541.