Many Objective Robust Decision-Making Model for Agriculture Decisions

Xavier Ignacio González – Guillermo Podestá – Federico Bert

Abstract

Farmers must yearly allocate fields to different crops management options. These decisions are critical because they cause a large spectrum of consequences. To support farmers decisions, this study introduces a framework based on Many Objective Robust Decision-Making methodology. The model has following features with respect to traditional decision support models: it shapes deep uncertainty about economic and climatic conditions by a scenario approach, instead of representing the context variables by single probability distributions; it considers a total of seven objectives that focus on initial costs, earning, losses, and returns and margin expected values to integrate a comprehensive range of farmers goals; the framework suggests robust strategies to farmers, those that pay back acceptable outcomes for as many scenarios as possible, rather than finding an 'optimal' strategy that optimizes one or several objectives; it identifies key scenario factors that affect the result by a supervised machine learning decision tree classification algorithm; and it provides visualizations that unveils alternative strategies that mitigate the bad performances identified. The framework is demonstrated with a numerical study case for a farm in the Pampas, where some strategies, as mix of crop alternatives involving soybean, wheat and maize are assessed and some of their vulnerabilities are identified and analyzed. All scripting and data are shared to the community for reproducibility.

Keywords: Agriculture – Crop Plan Decisions – Multi Objective – Robust Decision Making – Decision Making – Scenario Planning.

1. Introduction

Practiced at a global scale, agriculture is a primary human activity that has the challenge of producing food for billions of people while, at the same time, must deal with dynamic environmental, economic and cultural contexts. Climate change, market variations, and regulation amendments for more sustainable resource planning compels farmers to constantly adopt new and innovative farm policies (Godfray et al., 2010; Oram, 1989). One major way in which farmers can adapt to a shifting context (e.g., climate or market conditions) is by changing the allocation of land among a set of different crops and agronomic managements: this is analogous to rebalancing a stock portfolio as markets change. The selection of crops to be grown, their acreage and their allocation within the farmland are the main land-use decision in farming systems. Cropping plan decisions are the result of a decision-making process where farmers weight up various objectives and constraints, usually supported by technical advisors (Nevo et al., 1994). Models of cropping plan have been largely studied in the literature because land use decisions have a large spectrum of consequences. The aim of modelling is to support farmers to use resources more efficiently, to provide policy makers a tool to assess landscape changes and design better policy, and to conduct research into environmental, agronomic, and economic topics (Jérôme Dury et al., 2012).

While uncertainty about climatic and economic context is, unquestionably, a determinant element of cropping plans, it is largely neglected in cropping plan decision models. When considered, context uncertainty (e.g. yields and prices) is represented by single probability distributions and models return the optimal possible strategy (i.e. acreage of crops), optimizing the expected value of an economic outcome (e.g. maximizing expected utility) (Dittrich et al., 2017; Jérôme Dury et al., 2012; Meyer, 2002). Although this approach has a strong tradition in agricultural economics and has proven extraordinarily useful, if often has been dismissed as unrealistic, as it does not represent accurately farmers' decision mechanisms. Under this paradigm, adaptation is limited by the uncertainties and imprecisions that afflicts prediction. This predict-then-act approach has been used in numerous applications, often with great success.

However, climate change violates the postulates of predict-then-act as it is associated with conditions of deep uncertainty, where decision-makers do not know or cannot agree on the prior probability distributions for inputs to the system model, and the value system used to rank alternatives (R. Lempert et al., 2004; R. J. Lempert et al., 2003, 2006).

This paper contributes with a Crop Plan Decision framework that implements an adaptation of a Many Objective Robust Decision-Making (MORDM) approach (Kasprzyk et al., 2013) to examine the performance of adaptation strategies (i.e. crop mix allocation) by many objective metrics calculated over a wide range of plausible futures driven by uncertainty about the future state of climate and socio-economic drivers. In this way, strategies that perform sufficiently well across a range of alternative futures can be identified and assessed − even without accurate and precise predictions of future climate (Adger et al., 2009). To achieve this, the framework includes the characterization of uncertainty with multiple views of the future or states of the world. It can also incorporate probabilistic information but rejects the view that a single joint probability distribution represents the best description of a deeply uncertain future. Rather, this framework uses ranges, or more formally sets, of plausible probability distributions to describe deep uncertainty. Furthermore, the framework offers some additional enhancements with respect to traditional methods for land allocation decision making. First, it considers many objectives explicitly, which can prevent farmers from inadvertently ignoring aspects of the problem (such as key planning objectives) (Kasprzyk et al., 2013). Those are seven diverse economic objectives that shape a wide range of farmers goals and weigh differently the aspects of the result, for instance: earning, losses, and costs. Second, the framework uses a robustness rather than an optimality criterion to assess alternative policies. There exist several definitions of robustness, but all incorporate some type of satisficing criterion. A robust strategy can be defined as one that performs reasonably well compared to the alternatives across a wide range of plausible future scenarios. The traditional subjective utility framework ranks alternative decision options contingent on the best estimate probability distributions. Such analysis generally suggests a single best or highest-ranking option. In contrast, this analysis suggests not a single robust strategy but a set of a few reasonable choices that decision makers can choose among. Lastly, this framework implements "scenario-discovery", which is a methodology that uses machine learning algorithms to find relevant cluster of cases in databases of simulated data. Conveniently interpreted as scenarios, these clusters help illuminate and quantify the tradeoffs among alternative strategies under deep uncertainty (R. J. Lempert et al., 2003, 2006; R. J. Lempert & Groves, 2010).

The paper organizes as follows. Next, Section 2 describes proposed crop plan decision model for a hypothetical farm. After a short introduction about the region, following sub-sections discusses the model elements: the scenario approach to shape the uncertainty, the arrangement of strategies that configures all feasible options for land allocation, how the combination of strategy-scenario is evaluated, and the set of seven objectives that offer different perspectives of evaluation. Following, Section 3 shows and discusses results obtained by inputting simulated data to the model. First, subsection 3.1 shares a data visualization of a multi-objective Pareto front analysis that unveils candidates for robust strategies. Then, sub-section 3.2 explains a trained decision tree that reveals vulnerabilities of one of the candidates, which are context conditions in which its performance does not meet farmer's goals. Focusing on those vulnerabilities, sub-section 3.3 examines two visualizations that explore alternative strategies that are able to mitigate the bad performances identified. Section 4 discuses briefly some conclusions and future lines of research. Finally, Appendix A details the formulation of the seven objectives considered.

2. Decision Model

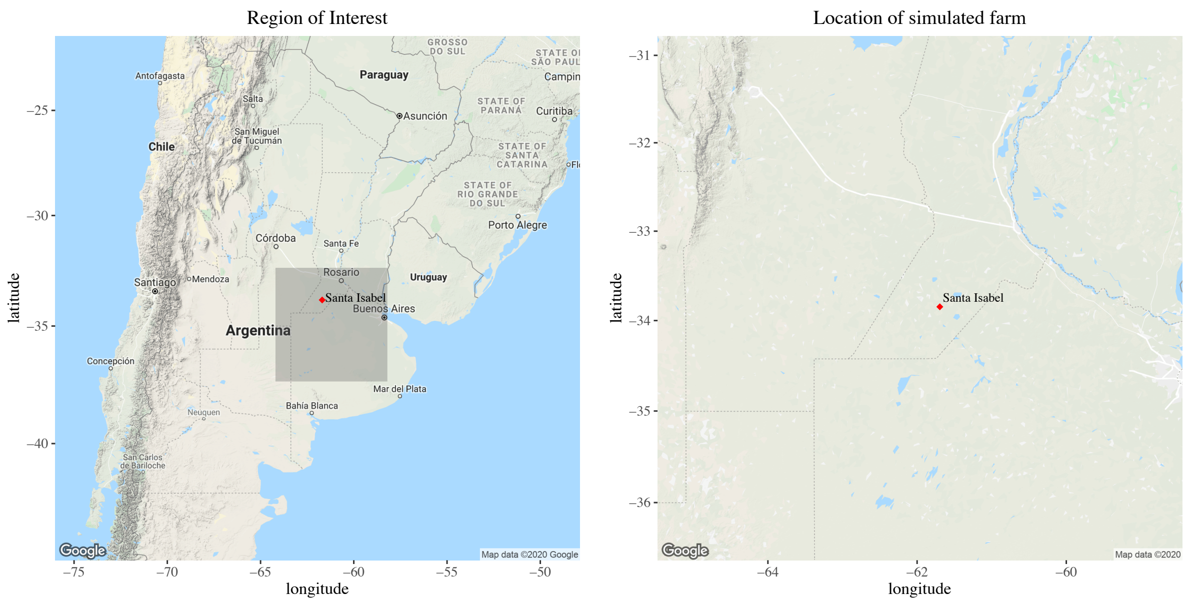

Figure 1. Location of a hypothetical farm where the crop plan decision is modeled.

The area targeted in this study is the region of central-eastern Argentina known as the Pampas, one of the main cereal and oilseed producing areas in the world (Calvino & Monzon, 2009). Crop rotations common in this region include maize, soybean, and a wheat. As the vast majority of field crops in the study area are rainfed, inter-annual climate variability (essentially fluctuations in precipitation) may induce considerable differences in crop yields from year to year. Within this area, a small village called Santa Isabel (33° 53' 19.2" S, 61° 41' 34.2" W) will be the reference site of a hypothetical farm where the crop plan decision model take place. This location, showed in the maps in Figure 1, is representative of a sub-region, called ‘core zone’ (*zona núcleo*), where agricultural activity is highly productive.

For this location, the decision to allocate farm's land among various activities (e.g., specific crops) was modeled. This decision-making process starts with the enumeration of all available and possible options for land allocation and each instance of this set is considered a different strategy. The selection of one strategy is made at the beginning of each cropping cycle (or agricultural year) and there is some, but not much, flexibility to revisit it later in the season. In consequence, land allocation must be decided in a context of uncertainty about upcoming conditions. The two major sources of risk to agricultural profits in the Pampas are climate variability and commodity prices drops. To characterize uncertainty within the cycle, multiple plausible views of the future are represented by scenarios.

2.1. Scenarios

Scenarios provide a commonly used and intuitively appealing means to communicate and characterize uncertainty in many decision support applications. The correct assembly of scenarios requires a thorough analysis by the modeler to contemplate a wide spectrum of situations; ideally, the set of scenarios should be both plausible and surprising (Xiang & Clarke, 2003). In RDM literature, scenarios are a set of future states of the world that enables, by exploration, the identification of vulnerabilities of proposed policies (i.e. crop allocations), that is, cases where a policy fails to meet its performance goals (R. J. Lempert & Groves, 2010). For this study, scenarios incorporate information about the climatic and crop price level contexts, which are the two main sources of uncertainty and were created based on historical data, for that was the data available to use.

The main variables that describe the climatic context of the crop cycle are related to the rainfalls and El Niño/Southern Oscillation (ENSO) phenomenon (Podestá et al., 2002). As there are virtually no artificial irrigation systems in the region, rainfalls are among the most important drivers of crop production performance. Although the region is considered homogeneously wet, rainfalls slightly vary from location to location. To estimate the rainfalls, modeled site was associated with the closest meteorological station. The climate records encompass the period 1961-2018 and include observed daily values of rainfall among other climatic daily variables. To depict climate conditions in terms that describe the weather at a seasonal level, daily rainfall measurements were summed up considering different time frames according to significant dates in the cropping cycle: ***pp_total***, from June 1st to June 1st next year; ***pp_jun2sep***, from June 1st to September 25th; ***pp_sep2dic***, from September 25th to December 1st; ***pp_dic2marNY***, from December 1st to March 30th of next year; ***pp_marNY2junNY***, from March 30th of next year to June 1st next year. Values for the amount of rainfall observed in each period are transformed to discrete numeric values, according to which quintile they belong, representing the level of rainfall: very low, low, medium, high and very high. This representation is more related to what farmers can use for their decision-making proposes. On the other hand, it is known that climate and crop yield variability is associated with the ENSO phenomenon. Therefore, a useful variable to describe the weather context in general terms is Oceanic Niño Index (ONI) for identifying El Niño (warm) and La Niña (cool) events [1]. A numerical variable ranging from -1.5 to 2 (i.e. strong La Niña to very-strong El Niño) was introduced to represent the cycle from June to June of next year. This variable, called ***ensoVar,*** completes the definition of a weather scenario.

A second set of variables that represents the uncertainty about the context is related to the commodity prices. Net from taxes, output prices vary significantly between the time when the crop plan decision is made and the moment the harvest is sold at the end of the cycle. Crop prices observations ranging from 1983-2018 and selected for the month of largest trading volume for

[1] https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

each crop were integrated in the scenario characterization. The historical price ranges (deflated to 2018 US dollars) are 84.5−286.5 \$ ton^{-1} for maize (***MT*** *and* ***Mt***), 184.9−472.1 \$ ton^{-1} for soybean (***S***), and 105.5−351.9 \$ ton^{-1} for wheat (***W***). Median prices for these crops are, respectively, 147.6, 305.9 and 167.0 \$ ton^{-1}.

Climate and crop prices are the main sources of uncertainty that affect the outcome in the crop plan decision. In this model, historical values of climate variables from 57 years and historical values of prices variables from 36 years were independently combined to generate 2052 plausible scenarios for each location modeled. For instance, one scenario is conformed with the triplets of market prices for soybean, wheat, and maize of the year 2001, and the rainfalls and ENSO event of year 1985. A summary of variables that describe each simulated scenario is detailed in Table 1.

	Scenario Variable ID	***Description***
Price variables	***W***	Wheat price
	S	Soybean price
	Mt	Early maize price
	MT	Late maize price
	S2W	Soybean to wheat price ratio
	W2M	Wheat to maize price ratio
	M2S	Maize to soybean price ratio
Climate variables	***QPP***	Quintile of accumulated rainfalls between June and June next year
	Qp_jun2sep	Quintile of accumulated rainfalls between June and September
	Qp_sep2dic	Quintile of accumulated rainfalls between September and December
	Qp_dic2marNY	Quintile of accumulated rainfalls between December and March next year
	Qp_marNY2junNY	Quintile of accumulated rainfalls between March next year and June next year
	enso	Oceanic Niño Index

Table 1. Summary of variables defining price and climatic scenarios.

Although the statistical independence between the climate variables and prices variables to be combined and generate scenarios may by arguably, analysis of data suggests so. The correlation matrix of variables corresponding to the same year in Table 2 indicates that: (i) there is a relatively low correlation, in absolute value, between climate and prices variables (maximum 0.38 between ***S*** and ***enso***), and (ii) there is a high correlation between variables within those two sets, for instance variable ***p_dic2marNY*** with ***pp_total***, and ***W*** with ***Mt***. Further considerations to capture the interaction between climate and prices variables will be matter of future research.

	S	***W***	***Mt***	***enso***	***PP***	***p_dic2marNY***
S	1.00	0.61	0.73	-0.05	-0.38	-0.36
W	0.61	1.00	0.79	-0.21	-0.14	-0.15
Mt	0.73	0.79	1.00	-0.22	-0.26	-0.28
enso	-0.05	-0.21	-0.22	1.00	0.25	0.01
PP	-0.38	-0.14	-0.26	0.25	1.00	0.59
p_dic2marNY	-0.36	-0.15	-0.28	0.01	0.59	1.00

Table 2. Correlation matrix of climate and prices observations in the same agricultural year.

Exploration of the set of scenarios or plausible future states (F) of the world facilitates the identification of robust strategies, (i.e. crop allocation mixes) that perform well for a wide range of scenarios, and their vulnerabilities referring to the states of the world where a crop allocation may fail to meet its performance goals as well as those where a strategy performance deviates significantly from the optimum policy in that state of the world.

2.2. Strategies

According to the formulation introduced in this study, a strategy refers to the decision involving the selection of the farmland usage at the beginning of the cropping cycle. The set of viable strategies in a location varies with climate and soil characteristics, despite a certain degree of homogeneity within the region. For the location modeled, six viable land uses or cropping alternatives (CAs) were defined in consultation with experts from the Asociación Argentina de Consorcios Regionales de Experimentación Agrícola (AACREA, www.aacrea.org.ar), a nonprofit farmers' group that entered a partnership with us for the present study. CAs are specified by the combination of (i) a crop (maize, full-cycle soybean and combined wheat-soybean) and (ii) a representative agronomic management, including cultivar/hybrid, planting date, planting density or row spacing, and fertilization level. The six CAs defined as viable options for a farm at the case location are detailed in Table 3 for illustrative purposes.

Cropping Alternative ID	***Crop***	***Genotype characteristic***	***Planting date***	***Planting density or row spacing***	***Nitrogen fertilizer ($kg\ ha^{-1}$)***
Mt	Maize	Early	Sep-25	7 plants m^{-2}	150
MT	Maize	Late	Nov-30	6 plants m^{-2}	130
Ws-S	Wheat	short intermediate	Jun-25		140
	Soybean	IV medium	Dec-5		
Wl-S	Wheat	long intermediate	Jun-1		140
	Soybean	IV medium	Dec-5		
Siii	Soybean	III long Cycle	Oct-20	30 cm between rows	0
Siv	Soybean	IV long cycle	Oct-20	30 cm between rows	0

Table 3. Definition of agronomic management for each of the six cropping alternatives considered

A strategy consists of the allocation of land to one or more of the CA's defined for the location. We assumed that the simulated farm is divided into ten equally-sized plots, and that the minimum unit of land assignment is one plot within a farm – that is, only one CA can be assigned to each farm plot. The number of possible land allocations or strategies N_{LA} can be calculated as

$$N_{LA} = \frac{(N_{CA} + p - 1)!}{p!\,(N_{CA} - 1)!} \quad (1)$$

where N_{CA} is the number of cropping alternatives and p is the number of plots in a farm. This equation can be read as the number N_{LA} of multisets of cardinality p with elements taken from a finite set of cardinality k. With six CAs and ten plots in the simulated farm, the number of possible land allocation strategies is 3003. This number goes up quickly as finer spatial granularity or additional CAs are considered. The set of strategies thus defined (St) and the set of scenarios or plausible future states of the world (F) obtained are combined with each other, generating what is called in RDM literature as 'futures ensemble' given by the cross-product $E = St \times F$ (R. J. Lempert et al., 2006), which requires a performance evaluation function of each strategy in each scenario.

2.3. Evaluation: outcomes of strategy-scenario combinations

A crop plan decision made by farmers consists in the selection of one strategy among the set of viable options, looking forward the fulfillment of multiple and competing objectives, most of them related to monetary goals, e.g. maximization of profit, as commonly acknowledge. Besides the variety of ways that monetary goals can be expressed in quantitative terms, other criteria seek efficient resource allocation, e.g. money or land, and sustainability, as pointed by experts from AACREA and some authors in literature (Jérôme Dury et al., 2012). Often criteria to select strategies are conflicting. For instance, some strategies may pay back good profits but are riskier, and some are less likely to return loses but their cost is higher. To consider simultaneously multiple conflicting objectives is one of the biggest challenges in crop plan decision models. This study proposes a diverse set of metrics and objectives to evaluate strategies individually and strategies in each scenario. A brief introduction to the metrics and objectives is given below while the formulation and a calculation detail is dscussed further in Appendix A.

The first proposed objective is the fixed cost (**FC**), which designates the monetary value that a farmer must allocate to start the production at the beginning of the agronomic cycle. It includes input costs (e.g., fertilizer, seed, field labors), land rental, and management costs. This money is spent, or put at disposal, immediately after the strategy is chosen, therefore, it is a value that does not depend on the scenario and does not carry uncertainty; it only varies on the CAs selected and their acreage. Farmers will seek to minimize this objective.

Secondly, we calculated an extensively used metric to measure farm performance, the farm-wide-net-margin (**FWNM**). Given a strategy selected, the FWNM can be calculated as the gross income minus fixed cost (FC) and variable cost, which includes harvest, transportation, and commercialization costs. The gross income is determined by the scenario that happens and it is calculated as the sum of crop price times the crop yield. We used process-level models to simulate yields for each CA in each "weather year". Models in the Decision Support System for Agrotechnology Transfer (DSSAT) package simulate crop growth and development as a function of daily weather, soil type, and crop genetic characteristics. The DSSAT models have been calibrated and validated in many production environments including the Pampas. Model inputs, such as soil parameters, including soil moisture and water and nitrogen content at the beginning of the cycle, were provided by local AACREA experts. Daily observations of climatic variables and crop prices values at the harvest were input to the model by the set of simulated scenarios.

To explore how FWNM varies depending on the selection of each CA among scenarios, Figure 2 shows a scatterplot with the mean and standard deviation of the FWNM and two overlapping histograms of the FWNM for Late Maize (MT) and combined crop Intermediate-

Cycle-Wheat followed by Soybean (Ws-S) considering 3 initial levels of water: low, medium and high. Charts show that combined crop Ws-S is the CA that returns, on average among scenarios, the maximum FWNM, marked with a vertical blue line in right charts; MT is the alternative with the lowest standard deviation of FWNM, as a measure of risk; and Soybean margins are somewhere in the middle, displaying medium mean and standard deviation, for initial water levels considered. The position of the points in the graphs gives us indications of the existence of some diversification (i.e. CA allocation mix) that reduces margin deviation, as a measure of risk, since CAs margins are not perfectly correlated.

Over the FC and the FWNM defined, we calculated two metrics that shape more accurately farmers decision preference: Utility (**U**) and Return on Investment (**ROI**). In this case, we assume the traditional definition of Utility, which means that it refers to the total satisfaction received by farmers from the monetary value, the FWNM, obtained as a result of their production. The ROI is a performance measure used to evaluate the efficiency of an investment. ROI tries to directly measure the amount of return on a particular investment, relative to the investment's cost. To calculate ROI, the FWNM is divided by the FC, representing the cost of the investment. The result is expressed as a percentage or a ratio. These two metrics, U and ROI, shape preferences for two different and common circumstances in crop plan decision making. Utility is appropriate when the farmer possesses a fixed amount of land and has some flexibility to allocate a variable amount of money to the production. Conversely, ROI is adequate in cases where the farmer has a fixed capital and intends to fully assign it into the production, which requires a versatility to rent more land for crop production. Both metrics are calculated for each strategy in each scenario to generate the 'futures ensemble'. In order to compare and evaluate strategies individually, an aggregation of values on simulated scenarios is needed.

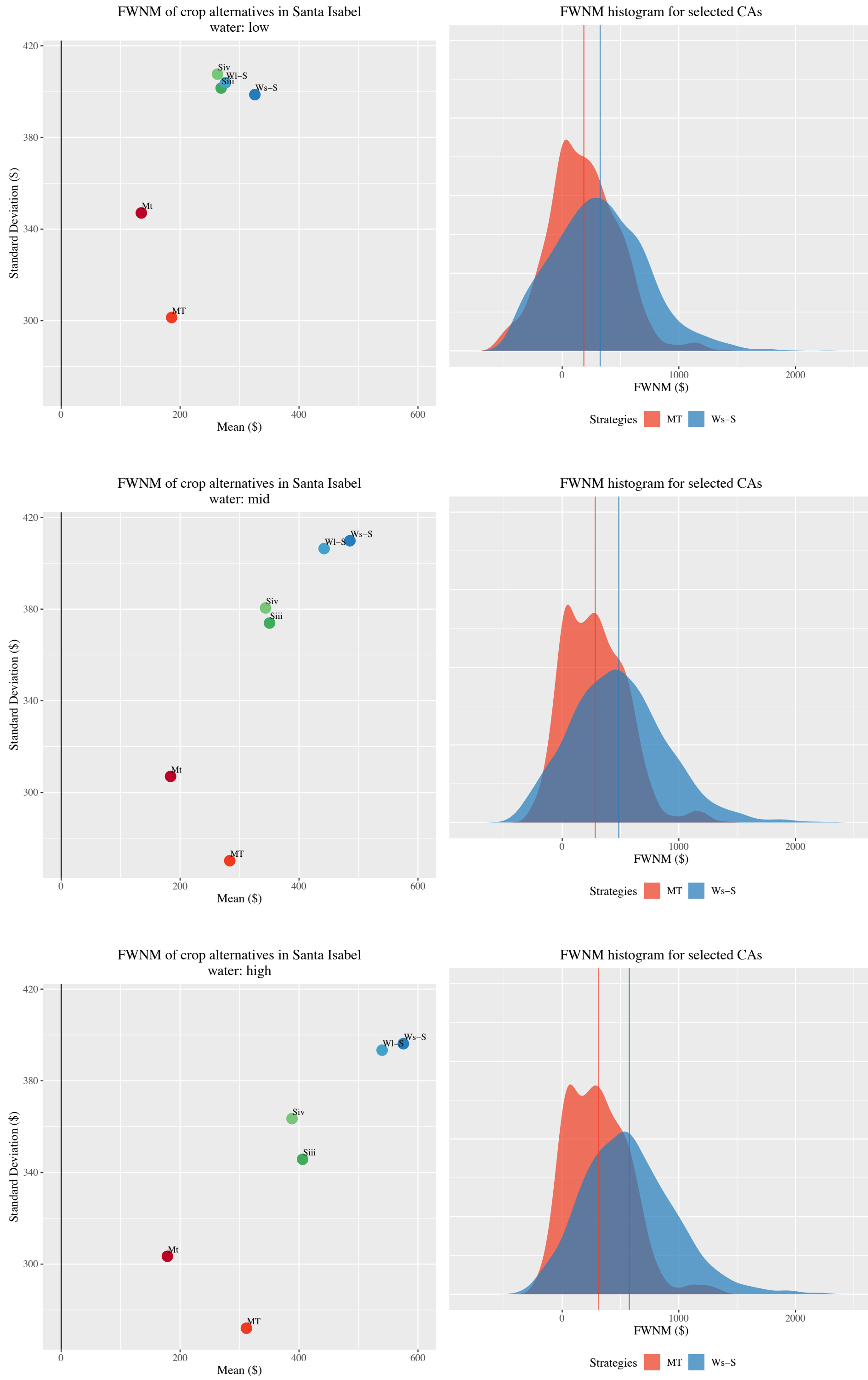

Figure 2. Mean and standard deviation of FWNM of 6 CAs and histogram of FWNM of CAs MT and Ws-S.

2.4. Evaluation: multi-objective assessment of strategies.

To evaluate the performance of the strategies over all simulated plausible scenarios, we define 3 ways to aggregate the Utility and ROI values. The first is to take their expected value considering equally probable scenarios.

Traditionally, expected utility (**E_U**) maximization is a predominant descriptive and normative model of choice under uncertainty in economics and a widely used criterion in agricultural decision making (Messina et al., 1999; Meyer, 2002). E_U hypothesis assumes that individuals evaluate uncertain prospects according to their risk attitude and expected level of satisfaction or “utility”. Despite its wide popularity in choice under risk modelling of agricultural applications, it does not entirely describe the diverse nature of the farmers’ goals, specially the way that farmers consider losses in the decision outcome. Analogously, we calculated the expected return on investment (**E_ROI**) as an alternative objective. This last metric takes into account the amount of money the farmer needs to have to start the production. ROI and U expected value calculations provide a good estimation of most likely outcome of selecting each strategy, but they do not consider how values are distributed in the distribution tails, that is, how possible occurrences are spread below and above the average, or, according to farmers’ perspective, between losses and gains.

The second aggregation defined to evaluate strategy outcome distributions is the conditional value at risk or expected shortfall (**ES**), a measure that estimates the risk of an investment in a conservative way, focusing on the less profitable outcomes. Specifically, "expected shortfall at q% level" is the expected return on the portfolio in the worst q% of cases (R. J. Lempert & Collins, 2007; McInerney et al., 2012). By selecting strategies with high value of ES, farmers seek to restrict the losses for the worst possible cases, ignoring what happens in other cases. This is a very conservative way to model the crop plan decision, although representative of actual decision drivers in the agriculture production. Again, we obtained for each strategy its expected shortfall of both: utility (**ES_U**) and return on investment (**ES_ROI**).

Thirdly, this work implements a widely-used measure in RDM literature that considers gains in relatively way (Cox, 2012; R. J. Lempert et al., 2006). The regret (**Reg**), or anticipated regret, is defined as the difference between the performance of a future strategy, given some value function, and that of what would have been the best performing strategy in that same future scenario. The regret of alternative strategies provides a conceptually and computationally convenient means to help identify robust land allocations and their vulnerabilities and also helps identify states of the world in which one strategy performs much better than another. Such information can help decision makers choose among strategies and improve their performance when they are unsure about the weighting over future states. Each strategy returns a regret value for each scenario considered, which is equal to the difference between the performance of the best possible strategy, the one selected having the 'crystal ball' for that scenario, and its own performance. The regret will be equal to zero in the event that the strategy is the best performing or will be greater than zero, in the case that it is not the optimal. To obtain a single value that measures the performance of each strategy incorporating results obtained for the set of scenarios, we calculate the 3rd quartile of the regret (**R3Q**), the limit that concentrates on its left 75% of the distribution of regret values. We apply the R3Q to utility metric (**R3Q_U**) and to the return on investment (**R3Q_ROI**). For example, suppose that a strategy results in a R3Q_ROI = 0.37, this means that, only in 25% of the scenarios, the difference between the best possible ROI and the ROI obtained is greater than 0.37; In 75% of cases, the difference with the best is smaller. Farmers will seek to minimize this objective.

In summary, we generated six objectives by combining three different types of aggregations -Expected Value, Expected Shortfall, and third quartile of Regret- with two complementary metrics -EU and ROI-. Those, plus the fixed cost (FC), constitute the set of seven objectives, listed in Table 4 and whose formulation is detailed in Appendix A, through which strategies are evaluated. These objectives model a diverse set of perspectives under which farmers assess their different options and allow different paths to explore the futures ensemble directed to find robust strategies.

ID	Description	Target	Focus on	Units
FC	Fixed cost	min	cost	[$]
ES_U	Expected Shortfall of Utility	max	losses	[$]
E_U	Expected Utility	max	average profit	[$]
R3Q_U	3rd quartile of regret of Utility	min	relative gains	[$]
ES_ROI	Expected Shortfall of ROI	max	losses	[%]
E_ROI	Expected ROI	max	average return	[%]
R3Q_ROI	3rd quartile of regret of ROI	min	relative earnings	[%]

Table 4. Summary of objectives considered

3. Results

This section describes an exploration of the futures ensemble generated by combining the 3003 strategies with the 2052 scenarios under the different objectives proposed for a particular hypothetical farm located in Santa Isabel, with a medium level water content at the beginning of the crop cycle. The exploration begins with a Pareto front analysis to discard strategies that would not make sense to choose and thus reduce the complexity. Then, data visualizations offer some insights about which strategies to consider as initial candidates for robust strategies. With an initial candidate as input, a machine learning algorithm trained on simulated data will unveil scenario conditions that are unfavorable to the candidate previously selected. The search for robust strategies alternatives to the candidate will focus on these context conditions identified.

3.1. Data Analysis: Pareto Front and visualizations

The seven objectives calculated for the set of strategies indicate how preferable strategies are in relation to others. For example, looking at Figure 2, for mid water content, it is expected that richer strategies in Ws-S return a high value of E_U, as its margin mean is the highest; on the other hand, strategies with more MT content are expected to return better values of ES_U, as the left tail MT margin distribution is shorter. Previous example shows how, within the set of all possible strategies, some combinations of CAs return better values for some objectives and some strategies return better values for other objectives. There are also some combinations that are surpassed by others in all the objectives considered. These strategies, according to this model, would make no sense to choose. To identify which strategies to consider and those to discard and reduce the complexity, we perform a Pareto Front analysis, in which we can separate strategies into two sets: those non-dominated and those dominated. For each dominated strategy we can find a combination of other strategies that overperforms in all considered objectives. We discarded dominated strategies using *Evolutionary Multi Objective Algorithm* R package (*emoa*).

Considering only the non-dominated set, a line plot in the higher chart of **Error! Reference source not found.** shows strategies as colored lines, and the y-axis marks the performance by each objective in the x-axis, where, the closer the line approaches the top, the better the performance; when the goal objective is the minimum, for instance R3Q_U, lower values are placed higher. Numbers within the cart indicate objective level. Line colors are obtained by mixing red, green, and blue for proportions assigned to maize, soybean, and wheat-

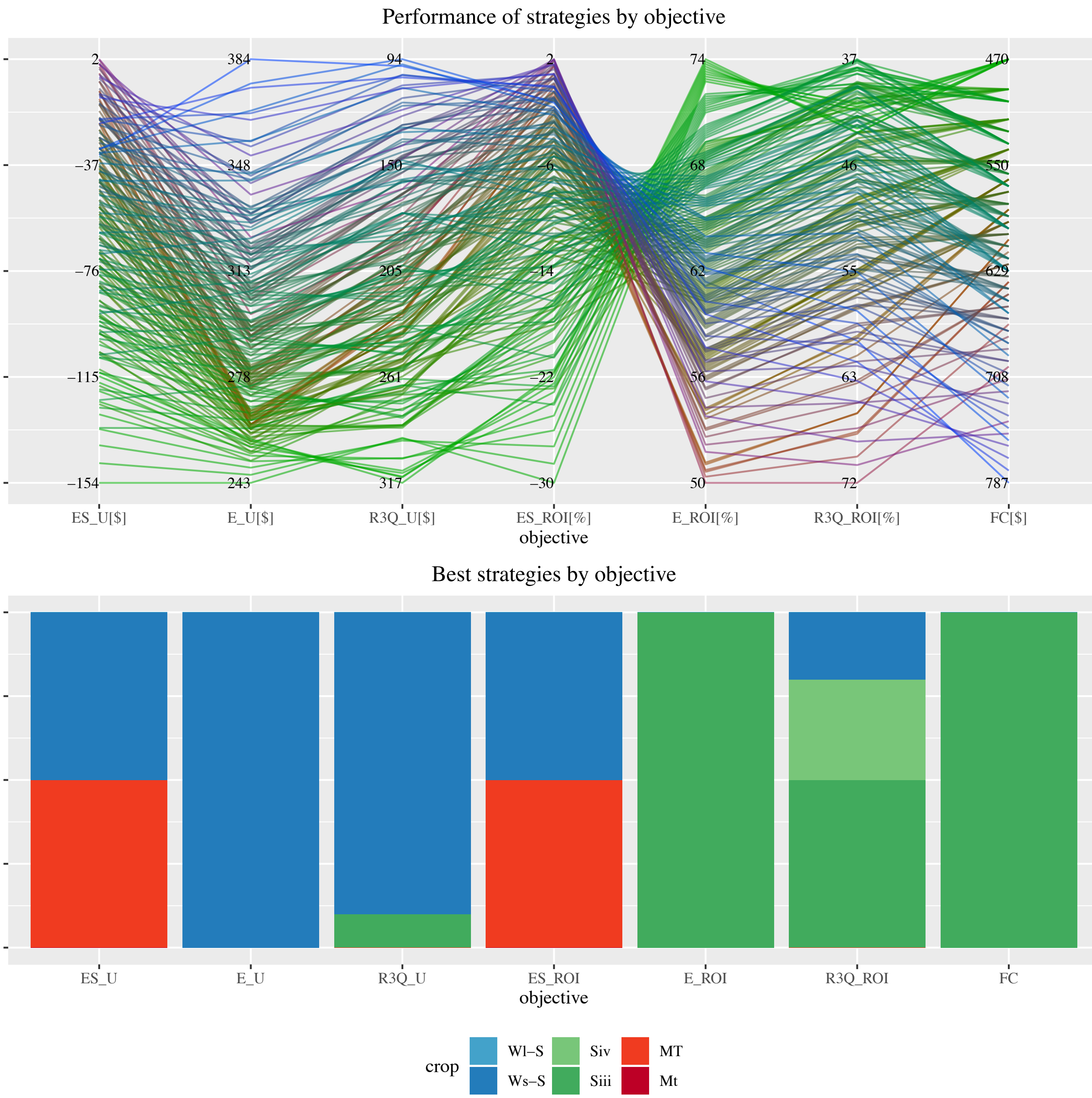

Figure 3. Performance of strategies and best strategies by objective

soybean, respectively. For instance, a farmland assigned to 50% wheat-soybean and 50% maize is drawn in violet, as a mix of red and blue, and one assigned to 1/3 of each crop is colored in brownish. Lower figure shows the crop mix allocation in a stacked bar plot for the best optimal solution for each objective considered.

Plots in **Error! Reference source not found.** give us evidence of following statements: (i) crop alternative S_{iii}, long-cycle soybeans, in green, is the strategy that requires less money to carry out, therefore it is an interesting alternative when farmer is looking to optimize fixed costs and ROI; (ii) crop alternative Ws-S, short cycle of wheat followed by soybean, in blue, is preferable according to objectives E_U and R3Q_U; (iii) mixing 50% of Ws-S and 50% MT, late corn, in red, is the strategy that best restricts the losses of metrics utility and ROI. Among variations of these 3 alternatives a proper decision should lie. Moving from one alternative to another entails a level of detriment in some objectives and of benefit in some others. Some of this compromise between objectives can be pictured within the colored region formed by the lines in the upper chart. To analyze further the trade-off and the relation between the context conditions and the objective performance, we applied machine learning to learn from simulated data.

3.2. Exploration by a Classification Tree

The first step in the robust decision-making process is the selection of one or more strategies as initial candidates for robust land allocation. In some cases, the candidate strategy may be suggested by decision-makers, for example, the recommended crop rotation for a given location. In this study, we initially select an allocation of 100% S_{iii}, long cycle soybean, the optimal strategy for the E_ROI objective, as a robust strategy candidate. Both the strategy and the objective are usual considerations by farmers in the area studied. A machine learning model applied to the futures ensemble will allow obtaining information on climatic and economic conditions in which the candidate strategy does not return satisfactory outcomes for proposed objective.

Various machine learning algorithms can be used to identify combinations of uncertain model input parameters that are most strongly predictive of decision-relevant regions in a database of simulation results (R. J. Lempert et al., 2008). For example, (Bryant & Lempert, 2010) used the Patient Rule Induction Method (PRIM) for this purpose. As an alternative to PRIM, here we build a classification tree to discriminate situations where initial candidate strategy has 'GOOD' and 'BAD' performance. The use of classification trees as an alternative to PRIM was explored in (R. J. Lempert et al., 2008). An advantage of the tree approach is that the sets of conditions that lead to a strategy's bad performance can be easily interpreted and communicated to decision-makers. The estimated classification tree identifies splits of input variables (i.e., portions of the hyper-rectangular, multivariate input space) that lead to either 'BAD' or 'GOOD' performance of candidate strategy: performance is considered 'GOOD' in scenarios for which ROI is $\leqslant$ 9.3% (the 25th percentile of ROI values for alternative S_{iii} under all scenarios); otherwise performance is 'BAD'. By definition, 1539 (or three-quarters) of the 2052 scenarios will show 'GOOD' performance; correspondingly, 513 scenarios will have 'BAD' performance. The predictors used to build a classification tree include all features described in Table 1. The tree algorithm successively divides the input space − one variable split at a time − with the goal of creating multiple regions that contain outputs of a single class. The tree is grown using 10-fold cross-validation to avoid overfitting. Each identified region corresponds to a terminal node (labeled "BAD" or "GOOD") of the tree in Figure 4.

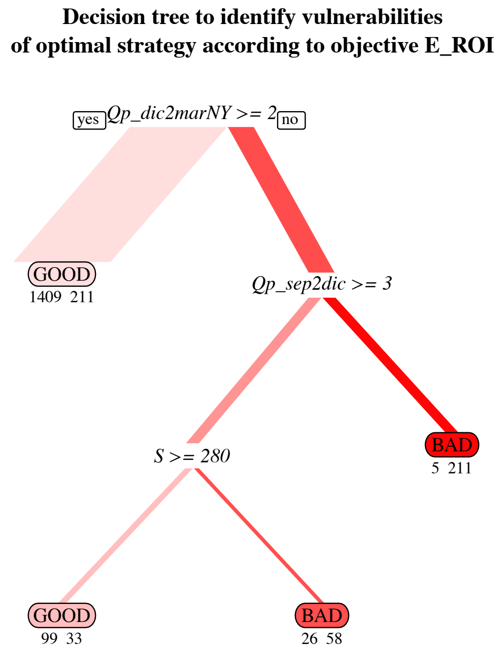

Figure 4. Classification tree predicting 'GOOD' or 'BAD' performance for land allocation strategy S_{iii}. Numbers in each terminal node indicate respectively the number of 'GOOD' and 'BAD' records in that node (note that the terminal nodes are not 100% of one class or the other, i.e., they are impure).

The first split in the classification tree describes the most relevant condition that defines scenarios in which alternative S_{iii} has a 'GOOD' performance. In this case, scenarios in which December-March rainfall totals above the lowest quintile limit (i.e., Qp_dic2marNY $\geqslant$ 2, or rainfalls > 338.6 mm) of the historical distribution produce an outcome where candidate strategy has a 'GOOD' performance. This period of time, December to March, in the southern hemisphere, corresponds to the reproductive stage of soybean development. As the median rainfalls in that period for all climatic scenarios is 496.5 mm, a value < 338.6 indicates a relatively unfavorable climate for soybean growing. Continuing along the right side of the tree (Qp_dic2marNY = 1), the successive branches indicate conditions that mainly lead to 'BAD' performance of candidate strategy. The second split (Qp_sep2dic $\geqslant$ 3, or Qp_sep2dic < 3 to the right branch) involves September-December rainfalls and points to a terminal node associated with a ‘BAD’ performance of candidate strategy. This can be understood as following. A scenario with a low level of rainfalls in both periods, first or second quintile from September to December, associated with the vegetative stages, and first quintile in the following period when the reproductive development stage take place, will lead to a ‘BAD’ result measured by the ROI for strategy S_{iii}. The last split of the tree (soybean price $\geqslant$ 280 \$ ton^{-1}, or soybean price < 280 \$ ton^{-1} to the right branch) indicates a ‘BAD’ performance in scenarios where soybean prices are relatively low, considering the median value of 306 \$ ton^{-1}, and rainfalls level in the reproductive stages is low. The tree obtained applying this methodology illustrates the vulnerabilities of a candidate strategy (in this case S_{iii}) as a crop option, that is, it shows the climatic and economic contexts that lead to an unfavorable result of the selection measured by one of the objectives (in this case the ROI). These contexts are easily interpretable, as they are defined by only 3-4 uncertain variables; interpretability drops as more variables are used to define the two performance regions. Further exploration can unveil alternative strategies that decision makers can follow to hedge risks associated with scenarios leading to bad performance of initial candidate strategies.

3.3. Identification of hedging strategies

The aim of this section is to identify alternative strategies that may perform nearly as well as strategy S_{iii} under the full ensemble of plausible scenarios but have better performance where S_{iii} is most vulnerable. As a first step, we use the classification tree grown in previous section (see Figure 4) to classify all 2052 scenarios into ‘GOOD’ or ‘BAD’ categories. That is, we divide the entire set of scenarios into two groups according to the predicted performance of initial candidate strategy in each scenario. To explore hedging alternatives, we evaluate the performance of all strategies for the two subsets of scenarios. For each strategy, we compute expected ROI values separately for ‘GOOD’ and ‘BAD’ scenarios. Then we use these quantities as coordinates for a scatterplot shown in Figure 5.

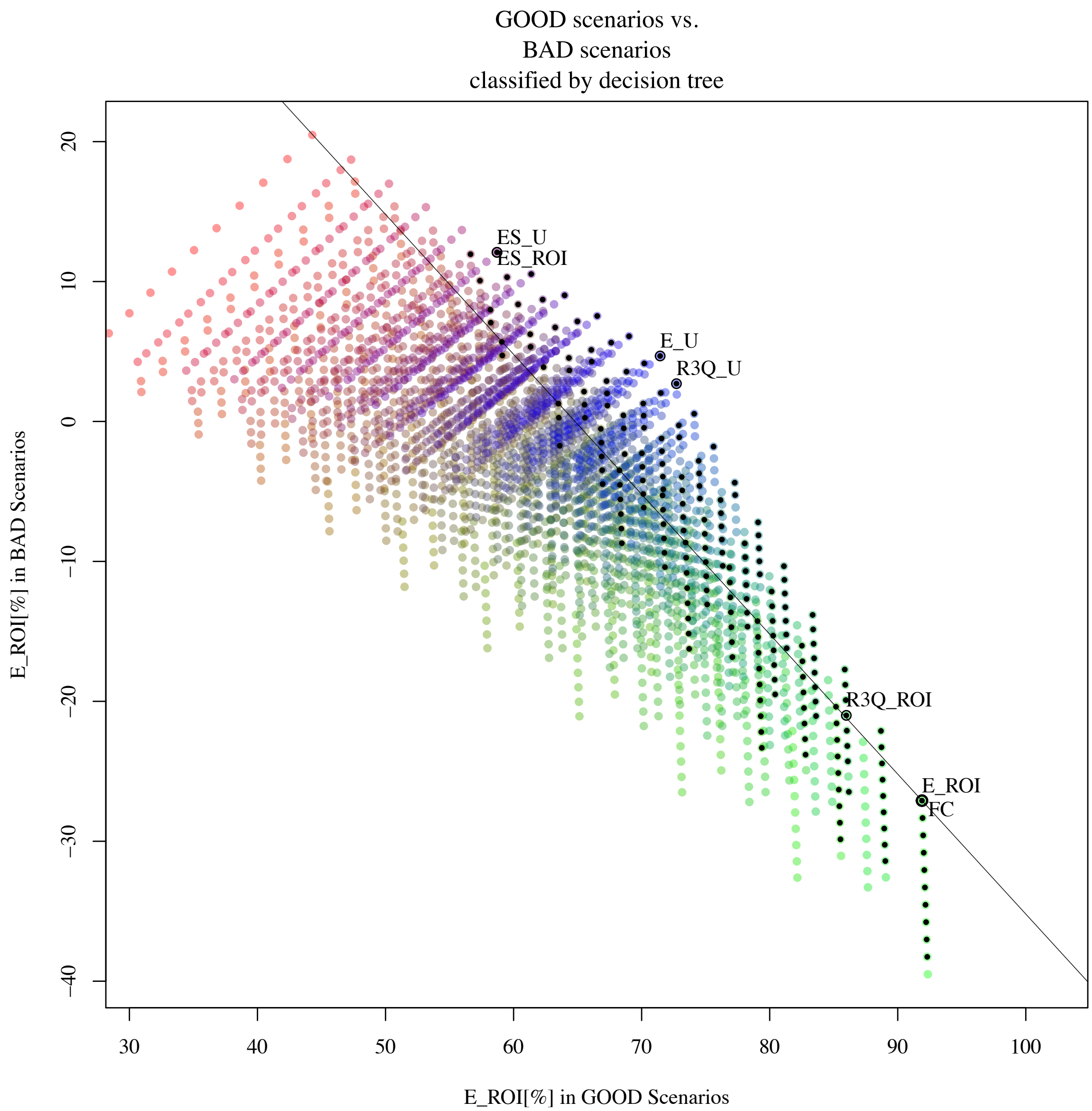

Figure 5. Trade-offs between strategies

Above chart illustrates E_ROI values of all strategies marked as points. Each color point represents a possible strategy farmer may choose, following the color definition described in the line plot in **Error! Reference source not found.**. Black dots are overlapped with colored points on those strategies that belong to the Pareto front and strategies that optimize considered objectives are labeled and marked with a small circle. Initial strategy S_{iii} is marked with labels 'E_ROI' and 'FCs', as it is the strategy that optimizes both objectives. After splitting scenarios into 'GOOD' and 'BAD', selecting the candidate strategy will result, in expected value, a ROI of 91.9% considering 'GOOD' scenarios, and a loss -27.1% in 'BAD' scenarios. Across this point, a diagonal line with slope -1 is traced, delimiting those points, located at the upper left, whose trade-off is favorable. For instance, strategy with label E_U, the one that maximizes the expected utility, offers a positive trade-off as it improves the outcome for 'BAD' scenarios in 31.8 percentual points (4.7% vs. -27.1% of E_ROI strategy) and worsen the result for 'GOOD' scenarios in 20.5% (71.4% vs. 91.9% of E_ROI), therefore it should be preferable according to this consideration, since the trade-off is favorable. Decision makers could choose strategies starting off the initial candidate and moving to the upper left over the diagonal line depending on their expectations of occurrence of 'GOOD' or 'BAD' scenarios and how much earnings and loses they are willing to give up and restrict, in 'GOOD' and 'BAD' scenarios respectively. This approach offers an exploration of E_ROI performance of all strategies among the two possible set of scenarios and answers with a another preference frontier, similar to the Pareto front, delimited by the three points marked with ES_U (and ES_ROI), E_U, and E_ROI.

Another perspective to explore the future ensemble is to compare the performance of two strategies among GOOD and BAD scenarios over the seven objectives all together, instead of considering the trade off on only one objective, like E_ROI in previous assessment. In this

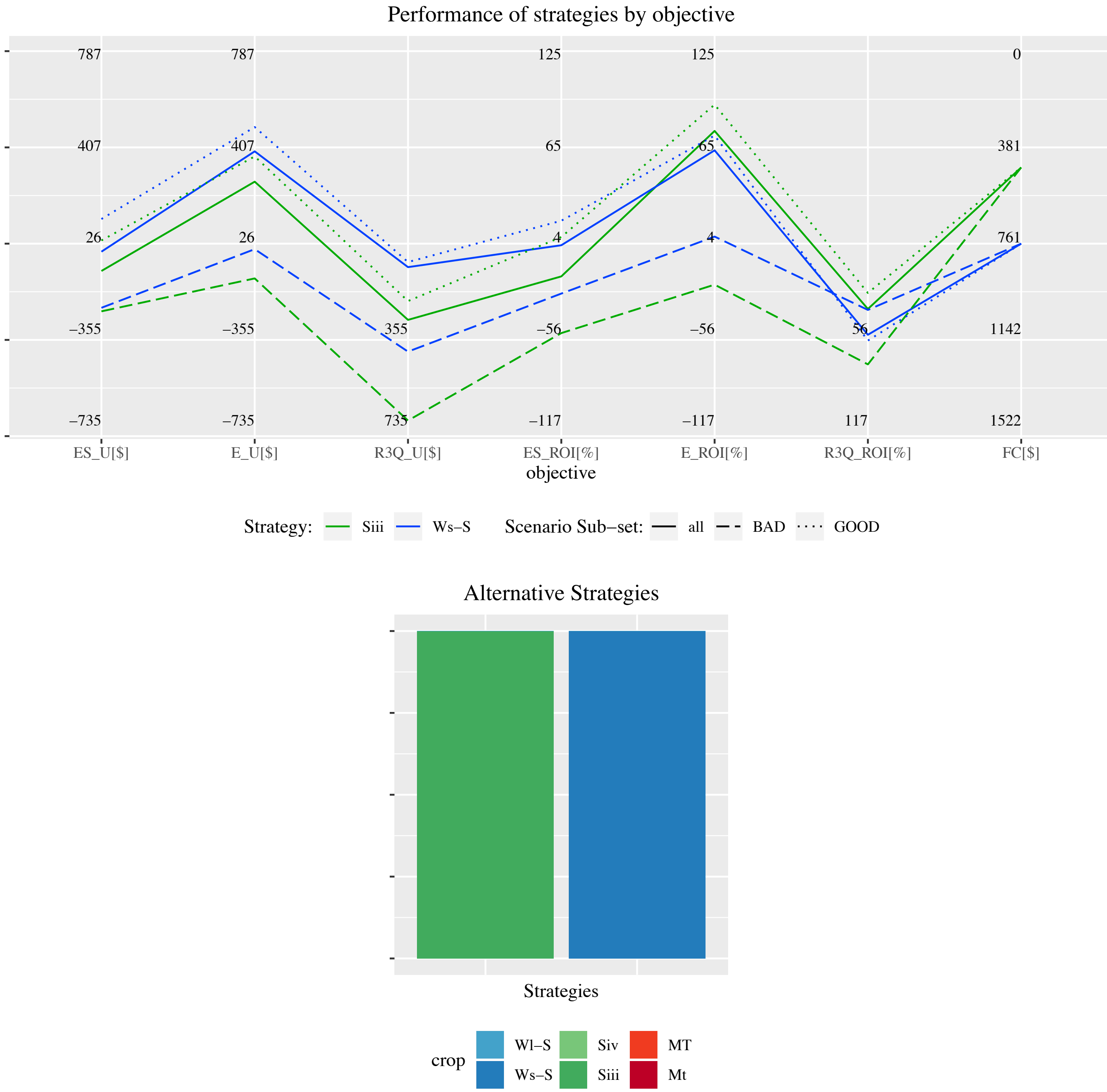

Figure 6. Line plot alternatives

analysis, we compare the initial candidate strategy, i.e. 100% of Siii, vs. the optimal strategy according to the E_U, i.e. 100% of Ws-S, as they appear to be strong candidates to robust strategy, nevertheless, any other two different points within the front can be explored as well. The top chart in Figure 6 compares in a line plot the outcome of the two candidates by the seven objectives defined for the total of scenarios and splitting the scenario set into GOOD and BAD subsets. Land allocation consisting in 100% combined Wheat-Soybean, Ws-S, outperforms crop alternative Siii, long-cycle Soybean, in four out of seven objectives when the entire set of scenarios is evaluated. Ws-S is superior, when no scenario restriction is applied, in all objectives involving Utility and also in the expected shortfall of ROI. Its alternative, Siii, is better in expected and third quartile of ROI and requires less fixed costs. This dominance persists when the evaluation considers only scenarios labeled as GOOD. On the other hand, when the assessment is restricted to scenarios labeled as BAD, crop Ws-S outperforms Siii in six out of seven objectives, all of them except of fixed costs. In addition, the difference in performances displayed as the distances between green and blue lines indicates an overall preference of Ws-S over Siii. Besides the disadvantageous trade-off on the expected ROI explained and depicted in Figure 4, there is another adverse trade-off regarding R3Q_ROI. Selecting strategy Ws-S over Siii will worsen in 31.5% the R3Q_ROI metric on GOOD scenarios, while on BAD scenarios, it will improve the metric in 35.9%, which represents a favorable trade-off. With this analysis and exploration, we found the vulnerabilities of initial candidate strategy, a land allocation of 100% long-cycle Soybean Siii, involving rainfall levels and Soybean price levels and obtained by the tree trained on simulated data. We analyzed how to mitigate those vulnerabilities by choosing an alternative strategy, crop Ws-S, and we confirmed that the mitigation is worth choosing when considering the trade-off involved regarding all objectives.

This procedure of defining a candidate, finding vulnerabilities, and identifying and assessing hedging strategies can be repeated many times with different initial strategies in order to obtain knowledge about the context condition that deteriorate performance of candidates and which crop alternatives can be produced to mitigate those situations. Final decision may be supported with additional sources of information such as: farmers experience, agricultural reports, quotations in derivates market, climate forecasts, etc.

4. Conclusions

We developed a framework to support yearly crop plan decision making to enable farmers deal with the deep uncertainty about the economic and climatic context. This framework models the uncertainty by a simulation of multiple scenarios (or plausible views of the future). Instead of employing scenarios traditionally to, for instance, structure the calculation of an evaluation metric expected value, in this framework they are explored by a classification algorithm with the aim of extracting relevant knowledge. This work used climatic and prices historical recorded data to generate scenarios and the 'probability' assigned to each scenario was equal. Nevertheless, other applications may include diverse ways of imagining and quantifying the set of scenarios to be explored and their parameters values not limited to historical records, for instance synthetic climatic time series (Verdin et al., 2019).

The main advantage of this framework over traditional methodologies is its capability to contribute not only which strategies are worth following, but also what are the 'dangers' that following each strategy deals towards the non-fulfillment of the objectives. Knowing the vulnerabilities, i.e. the key factors that affect the outcome, helps the decision maker to limit his interest to some sub-regions of the scenario space and thus enriches the understanding of the decision problem. Those sub regions of interest delimiting the vulnerabilities obtained will depend on, or are closed related to, the set of scenarios simulated a priori and the value range of its parameters. Focusing on those regions, the framework is capable to assess some alternatives that are able to mitigate the dangers found.

We demonstrated the model with a numerical study case for a hypothetical farm located within the Pampas in which we verified that traditional approaches may lead to different suggested strategies that do not satisfy accurately farmers' objectives. Starting with a candidate strategy of crop Siii assigned to the 100% of the land, a widespread crop practice within the region, the study found vulnerabilities related to low levels of rains between September and March, and low soybean prices. In those circumstances, a strategy corresponding to 100% of land assigned to Ws-S should be a better choice, and considering the trade-off between objectives, it should be a robust decision for farmers.

All code and data to reproduce the results of this paper are available at [2]. A new study that follows this one will provide with the explanation in detail of the data features and all functions used to model by the scenario-strategy approach, classify bad scenarios, calculate objectives and metrics, and visualize the results.

Another future line of research as a continuation of this work is the construction of a decision support system (DSS), web based and open to the public, in which farmers can input their personalized data related to costs, feasible crop managements and strategies, location, soil type etc. and the DSS return a subset of interesting alternatives to evaluate and the key context factors that harm the performance of each strategy identified.

Acknowledgements

This research work was partially supported by Peruilh PhD Scholarship of the School of Engineering of the University of Buenos Aires. The authors would like to thank www.prorindes.com and AACREA who provided most of the data used in this study.

[2] https://github.com/xavierign/MORDMAgro

References

Adger, W. N., Dessai, S., Goulden, M., Hulme, M., Lorenzoni, I., Nelson, D. R., Naess, L. O., Wolf, J., & Wreford, A. (2009). Are there social limits to adaptation to climate change? *Climatic Change*, *93*(3–4), 335–354. https://doi.org/10.1007/s10584-008-9520-z

Bert, F., North, M., Rovere, S., Tatara, E., Macal, C., & Podestá, G. (2015). Simulating agricultural land rental markets by combining agent-based models with traditional economics concepts: The case of the Argentine Pampas. *Environmental Modelling & Software*, *71*, 97–110.

Bryant, B. P., & Lempert, R. J. (2010). Thinking inside the box: a participatory, computer-assisted approach to scenario discovery. *Technological Forecasting and Social Change*, *77*(1), 34–49.

Calvino, P., & Monzon, J. (2009). Farming systems of Argentina: yield constraints and risk management. *Crop Physiology: Applications for Genetic Improvement and Agronomy*, *51*, 70.

Cox, L. A. (2012). Confronting Deep Uncertainties in Risk Analysis. *Risk Analysis*, *32*(10), 1607–1629. https://doi.org/10.1111/j.1539-6924.2012.01792.x

Dittrich, R., Wreford, A., Topp, C. F. E., Eory, V., & Moran, D. (2017). A guide towards climate change adaptation in the livestock sector: adaptation options and the role of robust decision-making tools for their economic appraisal. *Regional Environmental Change*, *17*(6), 1701–1712. https://doi.org/10.1007/s10113-017-1134-4

Dury, J, Garcia, F., Reynaud, a, Therond, O., & Bergez, J. E. (2010). Modelling the Complexity of the Cropping Plan Decision-making. *International Environmental Modelling and Software Society*, 8. http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

Dury, Jérôme, Schaller, N., Garcia, F., Reynaud, A., & Bergez, J. E. (2012). Models to support cropping plan and crop rotation decisions. A review. *Agronomy for Sustainable Development*, *32*(2), 567–580. https://doi.org/10.1007/s13593-011-0037-x

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., & Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. *Science*, *327*(5967), 812–818. https://doi.org/10.1126/science.1185383

Kasprzyk, J. R., Nataraj, S., Reed, P. M., & Lempert, R. J. (2013). Many objective robust decision making for complex environmental systems undergoing change. *Environmental Modelling and Software*, *42*, 55–71. https://doi.org/10.1016/j.envsoft.2012.12.007

Lempert, R. J., Bryant, B. P., & Bankes, S. C. (2008). Comparing algorithms for scenario discovery. *RAND, Santa Monica, CA*.

Lempert, R. J., & Collins, M. T. (2007). Managing the risk of uncertain threshold responses: Comparison of robust, optimum, and precautionary approaches. *Risk Analysis*, *27*(4), 1009–1026. https://doi.org/10.1111/j.1539-6924.2007.00940.x

Lempert, R. J., & Groves, D. G. (2010). Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west. *Technological Forecasting and Social Change*, *77*(6), 960–974. https://doi.org/10.1016/j.techfore.2010.04.007

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios. *Management Science*, *52*(4), 514–528. https://doi.org/10.1287/mnsc.1050.0472

Lempert, R. J., Popper, S. W., & Bankes, S. C. (2003). Shaping the Next One Hundred Years:

New Methods for Quantitative, Long-Term Policy Analysis. In *Rand*. https://doi.org/10.1016/j.techfore.2003.09.006

Lempert, R., Nakicenovic, N., Sarewitz, D., & Schlesinger, M. (2004). Characterizing climate-change uncertainties for decision-makers. *Climatic Change*, *65*(1–2), 1–9.

Letson, D., Laciana, C. E., Bert, F. E., Weber, E. U., Katz, R. W., Gonzalez, X. I., & Podestá, G. P. (2009). Value of perfect ENSO phase predictions for agriculture: Evaluating the impact of land tenure and decision objectives. *Climatic Change*, *97*(1). https://doi.org/10.1007/s10584-009-9600-8

McInerney, D., Lempert, R., & Keller, K. (2012). What are robust strategies in the face of uncertain climate threshold responses?: Robust climate strategies. *Climatic Change*, *112*(3–4), 547–568. https://doi.org/10.1007/s10584-011-0377-1

Messina, C. D., Hansen, J. W., & Hall, A. J. (1999). Land allocation conditioned on El Nino-Southern Oscillation phases in the Pampas of Argentina. *Agricultural Systems*, *60*(3), 197–212. https://doi.org/10.1016/S0308-521X(99)00032-3

Meyer, J. (2002). Expected utility as a paradigm for decision making in agriculture. In *A comprehensive assessment of the role of risk in US Agriculture* (pp. 3–19). Springer.

Nevo, A., Oad, R., & Podmore, T. H. (1994). An integrated expert system for optimal crop planning. *Agricultural Systems*, *45*(1), 73–92.

Oram, P. A. (1989). *Sensitivity of agricultural production to climatic change, an update*. IRRI.

Podestá, G., Letson, D., Messina, C., Royce, F., Ferreyra, R. A., Jones, J., Hansen, J., Llovet, I., Grondona, M., & O’Brien, J. J. (2002). Use of ENSO-related climate information in agricultural decision making in Argentina: A pilot experience. *Agricultural Systems*, *74*(3), 371–392. https://doi.org/10.1016/S0308-521X(02)00046-X

Pratt, J. W. (1964). *Risk aversion in the small and in the large, Econometrics 32, Jan*. April.

Verdin, A., Rajagopalan, B., Kleiber, W., Podestá, G., & Bert, F. (2019). BayGEN: A Bayesian Space-Time Stochastic Weather Generator. *Water Resources Research*, *55*(4), 2900–2915. https://doi.org/10.1029/2017WR022473

Xiang, W. N., & Clarke, K. C. (2003). The use of scenarios in land-use planning. *Environment and Planning B: Planning and Design*, *30*(6), 885–909. https://doi.org/10.1068/b2945

Appendix: Formulation of Objectives

This section details the mathematical calculation of metrics and objectives applied in this study. Table 5 introduces indexes and variables used in the formulation.

name	description	total indexes/variables	type	units
i	scenario	nF = 2052	index	
j	strategy	nS = 3003	index	
k	crop alternative	nCA = 6	index	
x_{jk}	crop mix allocation	6x3003	variable	[]
Y_{ik}	yield of crop alternative k in scenario i	6x2052	data	kg ha^{-1}
P_{ik}	price of crop alternative k in scenario i	6x2052	data	\$ kg^{-1}
VC_k	variable cost of CA k	6	data	[]
FC_k	fixed cost of CA k	6	data	$

Table 5. Indexes and variable used in equations.

The farm-wide net margin (FWNM) for a scenario i and a strategy j is the metric used as a main component of six out of the seven objectives introduced. It can be defined as

$$FWNM_{ij} = \sum_{k=1}^{6} x_{jk} Y_{ik}\, P_{ik}(1\ - VC_k) - FC_j \tag{2}$$

where x_{jk} is the proportion of farm allocated to cropping alternative *k* that defines strategy *j*. This variable is a component of the land allocation vector $x_j = (x_{j1}, \dots, \dots, x_{jk}, \dots, x_{j6})$ subject to the constraints $0 \leq x_{jk} \leq 1$ and $\sum_{k=1}^{6} x_{jk}$. The gross income is the sum of the multiplication of the allocation mix, yield, and price $\left(x_{jk} Y_{ik}\ P_{ik}\right)$ of each crop alternative produced. Production costs associated with each CA can be divided into fixed and variable costs. Variable direct costs VC_k, can be defined as a proportion of the income since they are proportional to CA's physical yield; this item includes harvest costs, marketing fees and grain transportation to market. Fixed costs, one of the objectives introduced in Section 2.2.4, in contrast, do not depend on a CA's yield; examples of the later include tillage and crop protection (but not harvest), seed and agrochemicals. The value of the fixed cost for a strategy *j*, FC_j, can be calculated as the sum of the fixed cost of the 6 crop alternatives weighted by the crop mix allocation.

$$FC = FC_j = \sum_{k=1}^{6} x_{jk} FC_k \tag{3}$$

On top of the *FC* and the *FWNM* defined, two metrics were introduced: Utility (U) and Return on Investment (ROI). The utility *U* refers to the total satisfaction received from an economic outcome and can be calculated with the formulation by (Pratt, 1964) where *r* is the risk aversion parameter and w_0 is the initial wealth of a farmer. The value of parameter *r* was set to 1.5, corresponding to a moderately risk averse farmer (Messina et al., 1999). The value of w_0 was set to 800 \$ ha^{-1}, an equivalent amount of money needed to carry out the CA whose fixed cost is the highest.

$$U_{ij} = \frac{\left(FWNM_{ij} + w_0\right)^{1-r}}{1-r} \tag{4}$$

The return on investment ROI is traditionally calculated in percentage as follows.

$$ROI_{ij} = \frac{FWNM_{ij}}{FC_j}100\% \tag{5}$$

To evaluate and compare strategies, metric values obtained for U and ROI in each strategy-scenario combination are aggregated among scenarios. First, we calculate the expected value considering each scenario as equally likely to occur with probability $\frac{1}{nF}$, an approach widely used in farm decision modeling (Letson et al., 2009; Messina et al., 1999).

$$E\_U_j = \frac{1}{nF}\sum_{i=1}^{nF} U_{ij} \tag{6}$$

$$E\_ROI_j = \frac{1}{nF}\sum_{i=1}^{nF} ROI_{ij} \tag{7}$$

The expected utility approach traditionally used in the literature condenses the observations to an average value and does not take into account the extreme values of the possible results or, what we might call, the tails of the distribution. Instead, another objective successfully used in RDM literature (Cox, 2012; McInerney et al., 2012), the expected shortfall, focuses on the worst cases. Also known as conditional value at risk (cVaR), it represents the expected value of the q% worst result and can be calculated as a conditional expectation of a particular metric. For example, the ES of the Utility is equal to the expected value of Utility, given that Utility is lower than the value that accumulates on its left the q% of the distribution.

$$ES\_U_j = E_i\left(U_{ij} \;/\; U_{ij} < Q_{q\%}\right) \tag{8}$$

Considering the scenarios equally likely to occur, the expected shortfall can be approximated to

$$ES\_U_j = \frac{1}{nF^{Qq\%}}\sum_{F^{Qq\%}} U_{ij} \tag{9}$$

Where $F^{Qq\%}$ denotes that the sum is only in those scenarios in which the value is lower than the percentile q and $nF^{Qq\%}$ is the number of scenarios in that subset, approximately equal to $nF.q$. Similarly, we can calculate the expected shortfall of the ROI

$$ES\_ROI_j = \frac{1}{nF^{Qq\%}} \sum_{F^{Qq\%}} ROI_{ij} \tag{10}$$

The regret, or expected regret, is the third approach by which we evaluated the strategies. It was widely used in RDM literature (Cox, 2012; Dittrich et al., 2017; R. J. Lempert et al., 2003) and its advantage is that it allows comparing the alternatives in a relative way, that is, with respect to selecting other alternatives. A regret metric can be calculated as the difference between the maximum simulated Utility (ROI) considering all possible strategies for scenario *i* − that is, the performance of the best possible strategy for this scenario − and the Utility (ROI) margin of strategy *j* in the same scenario *i*.

$$R\_U_{ij} = U_{i\,max} - U_{ij} \tag{11}$$

$$R\_ROI_{ij} = ROI_{i\,max} - ROI_{ij} \tag{12}$$

This metric can be understood as the difference between the actual result and the best possible result assuming that the farmer has the ‘crystal ball’ to see the future and select the best possible strategy. Higher values, i.e. difference from the best decision, indicate bad performances, because the decision ended in a result far away from the best. To aggregate those values over scenarios and be able to compare strategies, we take a percentile approach, similarly to the ES calculation. For regret, instead of calculating the expected value of the *q* lowest tail, we calculate the percentile that captures to its right the *q* worst cases. Different percentiles *q* can be used in the regret aggregation and in the expected shortfall calculation. Our choice was tied to the local farmers’ perception that about 2-3 years in a 10-year period are “bad” (i.e., show low profits). As farmers seek to maximize the ES and we used q = 25% percentile for its calculation. Conversely, as regret is a minimization objective, we aggregate its values with a percentile at q = 75%, which is equivalent to delimiting the 25% worst cases on its left tail. The third quartile (Q_3, or the 75th percentile) of the distribution of regret values for each strategy over all scenarios is named as

$$R3Q_{U_j} = perc\left(R_{U_{ij}}, q = 75\%\right) \tag{13}$$

and similarly, the third quartile of ROI

$$R3Q_{ROI_j} = perc\left(R_{ROI_{ij}}, q = 75\%\right) \tag{14}$$

One last note is that, since Utility is dimensionless, we express it in terms of certainty equivalent, to transform their values into meaningful economic units ($\$$ ha^{-1}) (Bert et al., 2015; J Dury et al., 2010). To do this, we inverse the Utility calculation given in equation [4], which leads to.

$$E\_U_j[\$] = \frac{1}{\left[E_{U_j}(1-r)\right]^{1-r}} - w_0 \tag{15}$$

Similarly, we transform ES_U and R3Q_U. This allows a better interpretability when comparing the objectives.