Multi-objective methods for crop insurance premiums: framework proposal and a case study in sugarcane

Roberto F. Silva[1], **Marcos R. Benso**[2], **Gabriela C. Gesualdo**[2], **Eduardo M. Mendiondo**[2],
Antônio M. Saraiva[3], **Patrícia A. A. Marques**[4], **Alexandre C. B. Delbem**[1]

[1]Institute for Mathematics and Computer Science - University of São Paulo
São Carlos, Brazil

[2]São Carlos School of Engineering - University of São Paulo
São Carlos, Brazil

[3]Polytechnic School - University of São Paulo
São Paulo, Brazil

[4]Luiz de Queiroz Colege of Agriculture - University of São Paulo
Piracicaba, Brazil

roberto.fray.silva@gmail.com, {marcosbenso,gabriela.gesualdo}@usp.br

emm@sc.usp.br,{saraiva,paamarques,acbd}@usp.br

Abstract. *The increase in extreme climate events due to climate change has resulted in crop losses, quality losses, environmental and social impacts in agricultural areas. Insurance against extreme events is a vital tool adaptation to deal with the impacts of those hazards. However, few works consider the optimization of different dimensions related to this tool. This work proposes a framework to use multi-objective optimization models to better design and evaluate crop insurance premiums and conducts a case study for sugarcane premiums at São Paulo state in 2010. The framework can be adopted for different crops, objectives, and models. The case study showed that around 20% of the policies evaluated were efficient solutions from the farmer's point of view.*

1. Introduction

Extreme climate events affect every region globally and are becoming more frequent and intense due to continuous human-induced climate change [IPCC 2021]. Particularly in South America, projections indicate an increase in the temperature of the hottest days at about two times the rate of global warming [IPCC 2021]. Consequently, events such as heatwaves, heavy precipitation, and agricultural and ecological droughts tend to become

increasingly common. Therefore, to reduce the future impacts of these extreme events or natural hazards, it is crucial to understand the spatial and temporal variability of the vulnerability of people and economic assets [Formetta and Feyen 2019].

There is additional complexity in extreme events vulnerability and uncertainty regarding developing economies due to their greater economic dependence on climate-sensitive primary activities, such as agriculture. According to [Kim et al. 2019] from 1983 to 2009, droughts caused 166 billion USD of losses in maize, soy, rice, and wheat production worldwide, affecting three-fourths of the global harvested area.

These losses can be minimized and rewarded through adaptation tools such as extreme event insurance. In this context, insurance prevents financial damage from turning into long-term economic damage [Seifert-Dähnn 2018]. Therefore, models that work on identifying and quantifying risk, estimating costs, and optimizing the premium are needed to enable risk transfer [Merz et al. 2010] and [Winsemius et al. 2013].

This research aims at proposing a theoretical framework for using multi-objective optimization (MOO) tools to evaluate crop insurance premiums considering multiple dimensions, such as the premium value, total insurable value (TIV), and coverage, among other essential aspects from the multi-dimensional model proposed [Navarro et al. 2021]. This framework encompasses the main stages of the data lifecycle and is model-agnostic. Then, a case study will be conducted for sugarcane crop insurance in 2010 at São Paulo state at the city level from the point of view of the sugarcane grower. The empirical attainment function (EAF) multi-objective model will be used for evaluating the efficient and non-efficient solutions for a bi-objective problem.

The two main research questions in this work are: (i) how can multi-objective optimization models be used for evaluating crop insurance premiums, considering multiple dimensions and objectives?; and (ii) what percentage of the crop insurance policies for sugarcane at São Paulo state cities could be considered efficient for the year of 2010, considering the maximization of the TIV per area and the minimization of the premium value per area? The main contributions of this work are: (i) proposing a theoretical framework for using MOO to evaluate crop insurance premiums that is model-agnostic; and (ii) evaluating the proposed framework on a case study with real data.

2. MOO problems on the agricultural domain

Crop insurance is an *ex-ante* risk management tool that can be used to mitigate the impacts of extreme climate events. In this context, the farmer pays a premium to an insurance company to be protected when crop yields are lower than a stipulated level [Bowers et al. 1997]. Nonetheless, the optimization of insurance contracts requires the development of methods that take into account multiple variables, as demonstrated by [Mohor and Mendiondo 2017]. Other works, such as [Fadhliani et al. 2019], incorporated participatory data to improve farmer's satisfaction and participation in insurance programs, improving coverage areas and farmer's adhesion in vulnerable rural areas.

Several important works in the literature evaluated the use of MOO on agricultural domains. Two examples are the ones by [Klein et al. 2013] and [Dunnett et al. 2018]. [Klein et al. 2013] proposed a system containing a MOO model and a self-organizing map to adapt agricultural land management considering the potential impacts of climate

change. Their main contributions were : (i) considering objectives related to farm productivity and environmental aspects; (ii) evaluating different climate scenarios and land uses; and (iii) evaluating the use of a generic crop model to estimate productivity.

[Dunnett et al. 2018] evaluated environmental and economic aspects related to agricultural production, but from a climate-smart strategy perspective. Those authors proposed and implemented a framework named Climate Smart Agricultural Prioritization (CSAP). Its significant contributions were: (i) considering different climate adaptation strategies for the different scenarios; (ii) considering socio-economic criteria; and (iii) considering both biophysical and agricultural emissions models on evaluating land use. However, that are also more specific MOO models for individual farming systems, such as the ones by [Groot et al. 2012] and [Jiang et al. 2020].

Lastly, this work evaluates the use of an important MOO model, the EAF. It was proposed by [Da Fonseca et al. 2001]. According to [López-Ibáñez et al. 2010], the EAF can be defined as the distribution of the different results of a stochastic local search algorithm applied to a multi-objective problem, considering maximization or minimization objectives. Different algorithms can be used for the local search. After several runs of the chosen algorithm with different random seeds, the model can identify several curves representing the relations between the dominated and non-dominated solutions. Additionally, an important use of the EAF model is to compare the behavior of different MOO models in a graphical format. For an in-depth description of the EAF model, we refer the reader to the works by [Da Fonseca et al. 2001] and [López-Ibáñez et al. 2010].

3. Methodology

The methodology used in this work was comprised of the following steps:

1. Framework proposal for the use of MOO for crop insurance premiums definition and evaluation, considering all the steps of the data lifecycle (as described by [Ball 2012]) and multiple dimensions (as proposed by [Navarro et al. 2021]);
2. Data gathering from sugarcane crop insurance premiums and coverage on the state of São Paulo for the year of 2010 from the official Government database, "Sistema de Subvenção Econômica ao Prêmio do Seguro Rural" ([SISSER 2021]);
3. Data processing, considering three main steps: (i) objective definition, generating the two objectives: TIV per area and premium per area; (ii) identifying and eliminating cities that presented missing data; and (iii) detecting outliers using the Boxplot method and eliminating them from the dataset;
4. Framework implementation on a bi-objective problem for sugarcane insurance in all cities at the São Paulo state at the year of 2010, using the EAF model through the EAF tools library [1]. The following objectives were considered: (i) maximizing the TIV per area, which is defined as the maximum value that the insurance company will pay for the farmer if the policy is used; and (ii) minimizing the premium value per area, which is the price that the farmer will pay for the insurance policy;
5. Evaluation of the optimal curve generated by the model and the different percentiles, considering: (i) the percentage of non-dominated points; (ii) the number of points that presented the worst strategies (according to the MOO model); and (iii) an overall analysis of the "efficiency" of the crop insurance premiums for sugarcane at São Paulo state in the year of 2010.

---

[1]https://mlopez-ibanez.github.io/eaf/

4. Proposed framework and its components

As described by [Navarro et al. 2021, Hajkowicz and Collins 2007], and [de Brito and Evers 2016], it is essential to consider MOO and multi-criteria decision-making to address the potential effects of extreme climate events in agricultural areas. Therefore, we propose a theoretical framework that considers multiple dimensions and objectives to improve crop insurance design and evaluation. It also considers the main steps of the data lifecycle [Ball 2012] and the standard workflow for machine learning models [LeCun et al. 2015]. It is composed of the following ten components:

1. **Data collection**, considering historical geographic data from official databases. The main objectives to be evaluated are related to the dimensions described by [Navarro et al. 2021] and to important criteria that are used on crop insurance premium design: coverage area, coverage production volume, TIV per area, the premium per area, government subsidies per area, probability of occurrence and potential impact of specific natural hazards, and willingness to pay for the insurance. Nevertheless, additional data sources can be considered;
2. **Objective definition**, generating the objectives that may be relevant for the specific use case. It is essential to observe the importance of the spatial and temporal dimensions for defining the objective functions;
3. **Data processing**, considering detection and handling outliers and missing data;
4. **Exploratory data analysis**, considering statistical analyses of the different criteria selected at the second component, to better understand their distributions and identify local patterns;
5. **Definition of the objective functions**, considering the multiple dimensions, and maximization and minimization objectives;
6. **MOO model selection**, considering their different aspects and characteristics, as explored in the works by [de Brito and Evers 2016, Hajkowicz and Collins 2007];
7. **Data standardization**, normalization, or transformation, if needed, to improve the model's overall results;
8. **MOO model implementation**, considering hyperparameters analysis and the objectives defined in step 3. It is also possible to conduct experiments with different combinations of the criteria and to compare different MOO models;
9. **Final model evaluation and analysis**, considering the following analyses: (i) identification of the optimal strategy; (ii) spatial analyses, considering dominant strategies for each region; and (iii) an in-depth analysis of the worst strategies identified. The main objective of this step is to extract the most important insights from the model's results. Additional analyses can be conducted if needed.
10. **Decision-making** based on the results of step 9. This may involve several actions, such as: redesigning the crop insurance policy, redesigning the crop insurance premium, among others.

5. Case study - Sugarcane crop insurance in 2010

This section contains the results of implementing the framework proposed in Section 4 on a case study of sugarcane insurance premiums at São Paulo state. First, we will present the results of the exploratory data analysis. Then, we will present and explore the results of the EAF model implementation.

Figure 1 illustrates the kernel density estimations (KDE) of the distributions for the premium per area (on the left, item "a") and the TIV per area (on the right, item "b"). It is essential to observe that: (i) both distributions are considerably different, indicating that the two objectives present a degree of conflict, justifying the multi-objective analysis; (ii) the normal distribution cannot be considered as a proxy for any of those distributions; (iii) the premium per area has a minimum value of R$13.39, a maximum of R$61.59, and a peak of around R$20.00; and (iv) the TIV per area has a minimum value of R$1180.62, a maximum of R$2530.18, and a peak of around R$2000.00. Among the main reasons that may explain those observations are: (i) the characteristics of the different regions, in terms of technology, soil, climate, and productivity;
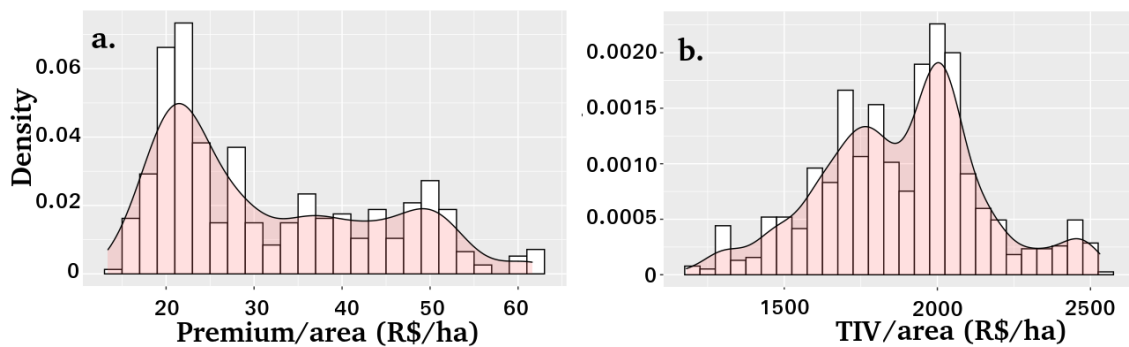

Figure 1. KDEs of the premium per area (on the left, item "a") and the TIV per area (on the right, item "b") objectives.

Figure 2 illustrates the results of the EAF model implementation, considering the bi-objective analysis described in Section 3. The EAF model is a stochastic optimization algorithm. Several simulations are run considering the data distributions and the two objective functions. The different colors in Figure 2 illustrate the areas that are related to the percentiles of the results obtained by those simulations. For example, the dark green area ("Best") represents the area that can be considered optimal for the MOO. The points that are on the upper border of the dark green area represent the efficient solutions found by the simulations conducted with the EAF model for the two objectives analyzed. The remainder of the points in that area (which are not along the upper border) represent solutions that can be considered good in practical terms, but that are not efficient (due to the fact that they are dominated by the solutions in the upper border).

The percentiles on the chart (represented by the different colors) represent the results of the run of the EAF model. Therefore, there may be combinations of the objectives that are not feasible in the real world. The efficient solutions (closer to the upper border of the green area on Figure 2) are candidates for Pareto-optimal solutions, based on the simulations that were conducted using the EAF model. These are more efficient from the farmer's point of view, as they will cost less and provide higher loss coverage.

It is vital to observe in Figure 2 that 162 points can be considered efficient or non-dominated. This indicates that around 20% of the total points in the dataset are non-dominated. Therefore, it is possible to conclude that around 20% of the total insurance policies in the dataset evaluated provide the best possible solution for the farmers. Additionally, around 30% of the policies in the dataset are encompassed in the dark green area ("Best"), meaning that they are also advantageous for the farmers.
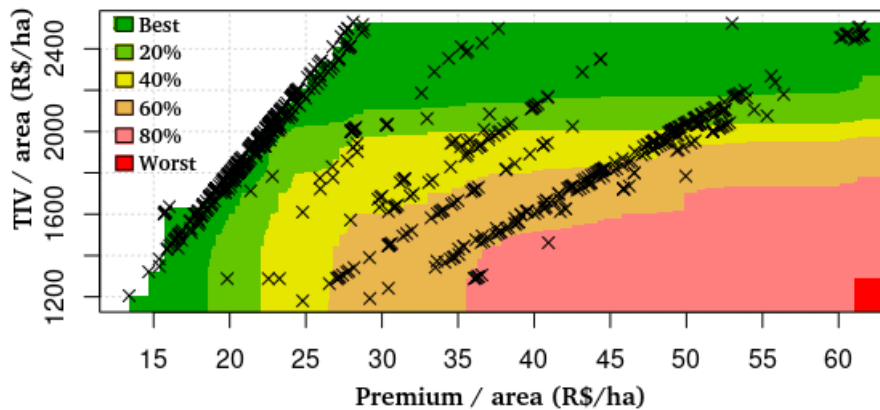
Figure 2. Final results of the EAF model implementation.

Lastly, no insurance policies are part of the dark red area ("Worst"), and very few insurance policies (7 data points in the evaluated dataset) are part of the 80% percentile (in light red). That means that very few policies are significantly inefficient for the farmers. Therefore, we can infer that most of the insurance premium prices for sugarcane in 2010 were aligned with the farmers' best interests. However, several insurance policies could be improved by reducing the premium per area or increasing the TIV per area (areas represented in medium and light green, orange, and light red).

6. Discussions

The results of the study case have shown that the use of multi-objective models could lead to interesting insights into the evaluation of crop insurance policies and premium rates. Therefore, the answer to the first research question is that, due to the several dimensions that may conflict with each other on designing and evaluating crop insurance policies, it is suitable to use an approach that considers partially or entirely conflicting objectives and decision criteria. The MOO models are specifically defined to deal with situations with conflicting objectives and decision criteria that must be maximized or minimized. As for the second research question, it was observed that 20% of the policies evaluated in the case study could be considered efficient from the farmer's point of view. It was also observed that an additional 30% of the policies were advantageous for the farmers.

Different supply chain stakeholders could use the results obtained from the case study for improving decision-making. Banks and insurance companies could focus their efforts on designing insurance policies with premium rates that increased the insurance coverage (in terms of area or production volume). Farmers could use the model's results to evaluate better and negotiate their insurance policies. Additionally, they could compare their policies with other policies in the same region. Lastly, it is essential to emphasize that an in-depth study considering different years, additional objectives (especially objectives related to the potential impacts of extreme climate events), and MOO models could improve the quality of the analysis. This could be used as an important tool for improving the decision-making processes related to public policies.

Several extensions of the case study conducted can be implemented, such as: considering additional areas, crops, objective functions, criteria, and MOO models. However, it is vital to observe that all extensions can be implemented following the framework pro-

posed in Section 4. Finally, the main limitations of this research were: (i) the lack of data available for implementing all the framework dimensions; (ii) the wide variety of options of MOO models; and (iii) the lack of applications of MOO for crop insurance premiums, which limited the comparisons that could be done with previous works in the literature. However, we believe that those limitations were correctly addressed in our work.

7. Conclusions and future works

A critical impact of climate change is the increase in frequency and impact of extreme climate events, such as droughts, floods, heatwaves, among others. These will increasingly impact agricultural areas and products, leading to crop losses, quality loss, and social and environmental impacts. One essential tool to reduce the impact of those natural hazards is extreme event insurance or crop insurance. Additionally, the use of MOO models can improve the quality of this tool by incorporating several dimensions into its design. Among those dimensions, it is important to consider: the probability of occurrence of the event, its potential impacts, the farmer's willingness to pay, among others.

This work proposes a theoretical framework for using MOO models to evaluate crop insurance premiums, considering multiple dimensions and objective functions. This is essential to address better the impacts of climate change on agroindustrial supply chains. The framework contains ten components and encompasses the whole data lifecycle. It also considers the main stages of the machine learning workflow, allowing the use of different machine learning models on the different components. A case study was conducted to evaluate sugarcane premiums at São Paulo state in 2010, concluding that around 20% of the policies were candidates for Pareto-optimal solutions, and that around 30% can be considered efficient policies from the farmer's point of view.

We recommend as future works: (i) implementing and analyzing other crops, objectives, and MOO models; (ii) and conducting a temporal analysis of the efficiency of the insurance premiums for a specific crop. To study the robustness of the EAF approach with TIV-Premium output, we recommend future works on: (i) attainment goal optimization with special attention to the slackness nature of some decision-making variables; (ii) setting the EAF simulations with TIV-Premium constraints. and (iii) scaling effects (upscaling/dowscaling) of spatial and temporal dimensions which strongly define both the objective functions and outputs.

References

Ball, A. (2012). *Review of data management lifecycle models*. University of Bath.

Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. (1997). *Actuarial Mathematics*. Society of Actuaries.

Da Fonseca, V. G., Fonseca, C. M., and Hall, A. O. (2001). Inferential performance assessment of stochastic optimisers and the attainment function. In *International Conference on Evolutionary Multi-Criterion Optimization*, pages 213–225. Springer.

de Brito, M. M. and Evers, M. (2016). Multi-criteria decision-making for flood risk management: a survey of the current state of the art. *Natural Hazards and Earth System Sciences*, 16(4):1019–1033.

Dunnett, A., Shirsath, P., Aggarwal, P., Thornton, P., Joshi, P. K., Pal, B. D., Khatri-Chhetri, A., and Ghosh, J. (2018). Multi-objective land use allocation modelling for prioritizing climate-smart agricultural interventions. *Ecological modelling*, 381:23–35.

Fadhliani, Z., Luckstead, J., and Wailes, E. J. (2019). The impacts of multiperil crop insurance on indonesian rice farmers and production. *Agricultural Economics*, 50(1):15–26.

Formetta, G. and Feyen, L. (2019). Empirical evidence of declining global vulnerability to climate-related hazards. *Global Environmental Change*, 57:101920.

Groot, J. C., Oomen, G. J., and Rossing, W. A. (2012). Multi-objective optimization and design of farming systems. *Agricultural Systems*, 110:63–77.

Hajkowicz, S. and Collins, K. (2007). A review of multiple criteria analysis for water resource planning and management. *Water resources management*, 21(9):1553–1566.

IPCC (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Cambridge University Press.

Jiang, S., Zhang, H., Cong, W., Liang, Z., Ren, Q., Wang, C., Zhang, F., and Jiao, X. (2020). Multi-objective optimization of smallholder apple production: Lessons from the bohai bay region. *Sustainability*, 12(16):6496.

Kim, W., Iizumi, T., and Nishimori, M. (2019). Global patterns of crop production losses associated with droughts from 1983 to 2009. *Journal of Applied Meteorology and Climatology*, 58(6):1233 – 1244.

Klein, T., Holzkämper, A., Calanca, P., Seppelt, R., and Fuhrer, J. (2013). Adapting agricultural land management to climate change: a regional multi-objective optimization approach. *Landscape ecology*, 28(10):2029–2047.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *N ature*, 521(7553):436–444.

López-Ibáñez, M., Paquete, L., and Stützle, T. (2010). Exploratory analysis of stochastic local search algorithms in biobjective optimization. In Bartz-Beielstein, T., Chiarandini, M., Paquete, L., and Preuss, M., editors, *Experimental Methods for the Analysis of Optimization Algorithms*, pages 209–222. Springer, Berlin, Germany.

Merz, B., Kreibich, H., Schwarze, R., and Thieken, A. (2010). Review article "Assessment of economic flood damage". *Natural Hazards and Earth System Sciences*, 10(8):1697–1724.

Mohor, G. S. and Mendiondo, E. M. (2017). Economic indicators of hydrologic drought insurance under water demand and climate change scenarios in a brazilian context. *Ecological Economics*, 140:66–78.

Navarro, F. A. R., Gesualdo, G. C., Ferreira, R. G., Rápalo, L. M. C., Benso, M. R., de Macedo, M. B., and Mendiondo, E. M. (2021). A novel multistage risk management applied to water-related disaster using diversity of measures: A theoretical approach. *Ecohydrology & Hydrobiology*.

Seifert-Dähnn, I. (2018). Insurance engagement in flood risk reduction – examples from household and business insurance in developed countries. *Natural Hazards and Earth System Sciences*, 18(9):2409–2429.

SISSER (2021). Apólices de seguros agrícolas - 2006 a 2016. *Available at: https://dados.gov.br/dataset/sisser3. Accessed on: 01.09.2021.*

Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., and Bouwman, A. (2013). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences*, 17(5):1871–1892.