Optimisation of Crop Configuration using NSGA-III with Categorical Genetic Operators

Extended Abstract

Oskar Marko
BioSense Institute
Novi Sad, Serbia
oskar.marko@biosense.rs

Dejan Pavlović
BioSense Institute
Novi Sad, Serbia
dejan.pavlovic@biosense.rs

Vladimir Crnojević
BioSense Institute
Novi Sad, Serbia
crnojevic@biosense.rs

Kalyanmoy Deb
Michigan State University
East Lansing, Michigan, USA
crnojevic@biosense.rs

ABSTRACT

One of the main tasks in agriculture is deciding which crop should be planted on which field. Agricultural companies often cultivate dozens of crops on hundreds of fields, making this problem extremely computationally complex. It was solved within evolutionary many-objective optimisation (EMO) framework. Objective functions included: profit, yield risk, price risk, scatteredness, crop rotation and environmental impact (total amounts of fertiliser and pesticide used). As the decision variables were categories (crops) and not real values, NSGA-III was adapted by changing the genetic operators of mutation and crossover from numerical to categorical. Optimisation was performed on the dataset provided by a partnering agricultural company. Out of the resulting population of solutions, characteristic crop configurations were chosen and compared to the benchmark, i.e. company's current strategy.

CCS CONCEPTS

• **Theory of computation → Mathematical optimization**.

KEYWORDS

EMO, genetic operators, optimization, NSGA-III, precision agriculture

ACM Reference Format:

Oskar Marko, Dejan Pavlović, Vladimir Crnojević, and Kalyanmoy Deb. 2019. Optimisation of Crop Configuration using NSGA-III with Categorical Genetic Operators: Extended Abstract. In *Genetic and Evolutionary Computation Conference Companion (GECCO '19 Companion), July 13–17, 2019, Prague, Czech Republic*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3319619.3321912

1 INTRODUCTION

FAO's estimates say that by the year 2050, we will need to increase the food production by 70 %, to feed the world's growing population [2]. For this reason, agriculture is shifting towards precision agriculture, where decisions are made on a high resolution, according to local specifics of crops, climate and the terrain. One of the main tasks that need to be optimised is concerned with crop configuration. Namely, every year a farmer or an agricultural company needs to decide which crop to plant on which field. This problem is combinatorial in its nature and its complexity grows exponentially with the number of fields. What is more, there are many confronting objectives among which the decision-maker needs to find the optimal trade-off.

GECCO '19 Companion, July 13–17, 2019, Prague, Czech Republic

ACM ISBN 978-1-4503-6748-6/19/07…$15.00
https://doi.org/10.1145/3319619.3321912

2 EXPERIMENT DETAILS

The study was conducted in cooperation with a large agricultural company that cultivates 5 different crops on 70 fields. The dataset consisted of 3 years of company's data and 8 years of crowdsourced data (802 samples in total). To estimate the profitability of each crop in the forthcoming season, probability density function (PDF) of crop yields in the next season was determined using statistical analysis of the historical data. Future prices were extrapolated from 20-year long time-series from the stock market and profit per crop was derived by multiplying the predicted yield and the predicted price. Yield risk was calculated as the variance of the PDF. It was a measure of yield instability, while the instability of produce prices was calculated as the variance in the time-series and added to the list of objectives as the price risk. The next objective was to decrease scatteredness of fields with the same crop. The corresponding objective function was calculated as the sum of fields' distances from the crop cluster centroid. Crop rotation is another factor that was considered and the cost of planting each crop, in relation to previously grown crop on the same field, was extracted from the dataset. The last two objectives were related to the environmental footprint of the crop configuration and were calculated as the total amount of fertilisers and pesticides used, respectively.

2.1 EMO with Categorical Genetic Operators

Due to its fast convergence and the ability to deal with many objectives, NSGA-III algorithm [3] was used in this study. Every company's field had one decision-variable associated to it, which indicated the crop that will be planted on that field. The number of

Table 1: Performance of characteristic solutions

	Profit	Price risk	Yield risk	Fertiliser	Pesticide	Crop rotation	Scatteredness
Max profit	28.80%	20.36%	-13.06%	-14.96%	-12.52%	-63.54%	-30.28%
Min risk	9.81%	9.30%	-9.08%	-6.95%	-3.24%	-36.06%	-21.31%
Sharpe	27.03%	14.49%	-7.16%	-11.25%	-13.31%	-10.89%	-24.19%

decision-variables was thus equal to the number of fields (70), while the number of objectives was 7, as explained in the previous section. The mutation probability was 1/70, while the crossover probability was 0.9. The set of categories for decision-variables included 5 most commonly grown crops in continental climates: wheat, maize, soybean, sugar beet and sunflower. In this setup, NSGA-III could not be used in its original form as it was designed for real-valued decision variables. It is, as many other EMO algorithms, based on the simulated binary crossover (SBX) [1] and polynomial PDF for mutation. The interpretation of these genetic operators is that a) the probability of offspring solutions is higher for solutions closer to either parent and b) that finer changes in a solution, due to the mutation, are more likely than the substantial ones. Rounding up the real-valued solutions would renounce these principles, as it would be susceptible to the ordering of categories and there is no single metric that could be used for ordering of crops on a common scale. The proposed solution is the following. Categories were suspended in a multi-dimensional space, where their coordinates were essentially the values of their objective functions. For categorical crossover, a line was drawn between the parent categories and the other ones were projected on it (Figure 1).

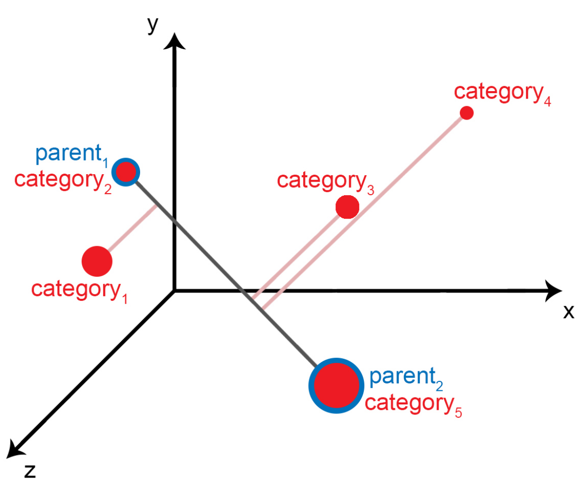

Figure 1: Categories suspended in a multi-dimensional space, where the coordinates were set as the values of objective functions.

The problem could now be observed as real-valued along this line and the usual PDF could be applied. As for mutation, categories were also suspended in the multi-objective space and now their distances from the parent category were observed. Polynomial PDF was applied, where the probability of each category was determined by its Euclidean distance from the parent category in the multi-dimensional space. In this way, there was a higher probability of children being closer to parents and one of the most important concepts of evolutionary computation was preserved.

3 RESULTS AND DISCUSSION

There were 20 runs of the code, each having 500 generations, with 100 solutions in each one. We focused on 3 most important strategies from the company's perspective: profit maximisation, cumulative risk minimisation and Sharp ratio (profit/risk) maximisation. They were compared to the benchmark (crop configuration proposed by the company's experts) and their relative performance is given in Table 1. Values of other objective functions serve as a proof that other stakeholders such as transport companies, government's environmental agencies or the society can benefit from such optimisation.

4 CONCLUSIONS

The topic of this study was crop configuration planning, one of agriculture's main problems that farmers are facing every year. It was mathematically defined as a multi-objective portfolio optimisation problem, where the crop configuration of choice should provide the optimal trade-off between 7 competing objectives. This was solved using NSGA-III, but due to the categorical nature of the decision variables (i.e. crops), genetic operators had to be adjusted. Categories were suspended in a multi-dimensional space and their Euclidean distance and projections were used in mutation and crossover. Optimisation yielded three characteristic strategies for the decision-maker to choose from according to his/her preferences. The results show that NSGA-III can be extremely useful in categorical optimisation problems and that there is a huge potential in the application of EMO in agriculture.

ACKNOWLEDGMENTS

The work was supported by the Ministry of Education, Science and Technological Development of Republic of Serbia under Grant No.: 44006 and Provincial Secretariat for Higher Education and Scientific Research of Vojvodina.

REFERENCES

[1] Ram Bhushan Agrawal, Kalyanmoy Deb, and Ram Bhushan Agrawal. 1995. Simulated binary crossover for continuous search space. *Complex systems* 9, 2 (1995), 115–148.
[2] Nikos Alexandratos, Jelle Bruinsma, et al. 2012. *World agriculture towards 2030/2050: the 2012 revision.* Technical Report. ESA Working paper FAO, Rome.
[3] Kalyanmoy Deb and Himanshu Jain. 2014. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE Transactions on Evolutionary Computation* 18, 4 (2014), 577–601.