Use of Data Analysis Techniques for Multi-Objective Optimization of Real Problems: Decision Making

António Gaspar-Cunha1, Francisco Monaco2, Paulo Costa1, Alexandre Delbem4

1 IPC-Institute for Polymers and Composites, University of Minho, Guimarães, Portugal; agc@dep.uminho.pt
2 Institute of Mathematics and Computer Science, University of São Paulo, Brazil; monaco@icmc.usp.br
4 Institute of Mathematics and Computer Science, University of São Paulo, Brazil; acbd@icmc.usp.br
* Correspondence: agc@dep.uminho.pt, University of Minho, Department of Polymer Engineering, Campus de Azurém, Guimarãres, Portugal

ABSTRACT

Most, if not all, real optimization problems can be seen as multi-objective since several objectives are to be satisfied concurrently and are often conflicting. Also, due to the high computation times usually required by the numerical modelling routines available to calculate the values of the objective function, as a function of the decision variables, it is necessary to develop alternative optimization methodologies able to reduce the number of solutions to be evaluated, i.e., if compared with the procedures typically employed, such as evolutionary algorithms. Moreover, in a multi-objective environment, it is also necessary at the end of the optimization process to select a single solution from the pool of optimal non-dominated solutions obtained.

Real industrial processes can be characterized by different types of data that can influence assertively its performance. For example, in the industrial process studied here, polymer processing, variables related to operating conditions of the machine, polymer properties and system geometry affect its operation since the thermomechanical environment developed allows obtaining mathematical relationships between these design variables and the objectives to be accomplished. This enables the direct process optimization using those routines to evaluate the solutions proposed by the optimization algorithms. However, this routine must be run several times, implying high computation times due to the sophistication of the numerical codes

This work aims to apply Artificial Intelligence based on a data analysis technique, designated by DAMICORE, to surpass those limitations, improve the optimization process and help the selection of the best-equilibrated solution at the end. An example from single screw polymer extrusion is used to illustrate the efficient use of a methodology proposed, with a focus on decision making.

Solving Multi-Objective Optimization Problems (MOOP) requires some interaction with a DM, for example, an expert in the field. The aim is to use data analysis techniques to reduce and improve the quality of those interactions, which can be done by integrating optimization methodologies with data analysis tools, i.e., the use of data to drive the optimization. At least, two different possibilities can be applied by data-driven optimization: i) replacement of the original method of calculating the objectives by a metamodel or surrogate, and 2) helping the computer in deciding about the best solutions to the problem.

The aim here is to use the DAMICORE framework to facilitate the optimization taking into account the limitations/characteristics referred to above. The DAMICORE structure is based on the estimation of distances by compression algorithms called Normalized Compression Distance (NCD). Then, a Feature Sensitivity Optimization based on Phylogram Analysis (FS-OPA) is used to find the set of principal features related to the real problem environment.

The present study focus on two levels of learning, which will be used to study an industrial case study using real data:

First-level learning – the aim is to find clusters of variables sharing information, designated by clades, each representing the set of variables with important interactions. The result of this level is a table with a list of variables with a cluster per row.

Second-level learning – the application of FS-OPA allows the estimation of the contribution of each clade of variables to the objectives, which is made by determining the distance between the clades of objectives (oclade) to each variable clade (vclade) using the phylogram obtained. These distances are an estimation of the influence of a clade to improve an objective. The results of this level are two different matrices, one with the phylogram distances from vclades to oclades and the second with the relative phylograms distances from each variable to each objective.

From a practical point of view, the application of this method to the data of each population of solutions previously obtained during the multi-objective optimization using evolutionary algorithms will allow capturing the interactions between the decision variables and the objectives and, in the end, select the most important objectives to the process. Therefore, the multi-dimensional space, that results from the six objectives existent in the problem solved, can be reduced, which will help the decision maker in selecting in an easy way the solution to be applied in real practice.

The results obtained for this practical example are in agreement with the expected thermomechanical behaviour of the process, which demonstrated that AI techniques can be useful in solving practical engineering problems.