Relevance of Multi-Objective Optimization in the Chemical Engineering Field

Geraldine Cáceres Sepúlveda

Submitted to the Faculty of Engineering

in partial fulfilment

of the requirements for the degree of

Master of Applied Sciences in Chemical Engineering

Department of Chemical and Biological Engineering

Faculty of Engineering

University of Ottawa

Abstract

The first objective of this research project is to carry out multi-objective optimization (MOO) for four simple chemical engineering processes to clearly demonstrate the wealth of information on a given process that can be obtained from the MOO instead of a single aggregate objective function. The four optimization case studies are the design of a PI controller, an SO_2 to SO_3 reactor, a distillation column and an acrolein reactor. Results that were obtained from these optimization case studies show the benefit of generating and using the Pareto domain to gain a deeper understanding of the underlying relationships between the various process variables and the different performance objectives.

In addition, an acrylic acid production plant model is developed in order to propose a methodology to solve multi-objective optimization for the two-reactor system model using artificial neural networks (ANNs) as metamodels, in an effort to reduce the computational time requirement that is usually very high when first-principles models are employed to approximate the Pareto domain. Once the metamodel was trained, the Pareto domain was circumscribed using a genetic algorithm and ranked with the Net Flow method (NFM). After the MOO was carry out with the ANN surrogate model, the optimization time was reduced by a factor of 15.5.

Résumé

Le premier objectif de ce travail de recherche est d'exécuter l'optimisation multi-critère pour quatre procédés simples découlant du domaine du génie chimique pour démontrer clairement la richesse des informations qui peuvent se trouver pour un procédé donné, pouvant être obtenues à partir de l'optimisation multi-objectif (OMO) au lieu d'utiliser une fonction objective unique agrégée. Les quatre systèmes étudiés pour l'optimisation sont la conception d'un contrôleur PI, la conception d'un réacteur de SO_2 à SO_3, la spécification du taux de reflux d'une colonne de distillation et la conception d'un réacteur de production d'acroléine. Les résultats obtenus à partir de ces cas d'étude montrent l'avantage de générer et d'utiliser le domaine Pareto pour mieux comprendre les relations sous-jacentes entre les différentes variables du procédé et les différents objectifs fixés.

En outre, un modèle d'une usine de production de l'acide acrylique est développé afin de proposer une méthodologie pour résoudre l'optimisation multi-critère du système à deux réacteurs utilisant des réseaux neuronaux artificiels comme métamodèles, dans un effort pour réduire l'exigence de temps computationnel qui est généralement très élevé lorsque des modèles phénoménologiques sont utilisés pour approximer le domaine de Pareto. Une fois que le métamodèle a été développé, le domaine de Pareto a été circonscrit à l'aide d'un algorithme génétique et classé selon la méthode de Net Flow. Après avoir effectué l'optimisation multi-critère en utilisant le métamodèle formé de réseaux neuronaux artificiels, le temps d'optimisation a été réduit par un facteur de 15.5.

Acknowledgments

First and foremost, I would like to express my deep gratitude to my supervisor, Dr. Jules Thibault who was the most patient and supportive professor one could ask during this research. I will always be grateful to him for sharing his knowledge with me and for helping me along the way. I would also like to express my sincere gratitude to Dr. Silvia Ochoa who co-supervised this work from Colombia and who gave me part of her invaluable time during this research. I would not have been able to complete this thesis without their continuous support, insights and invaluable guidance.

I would also want to thank my family and friends for their encouragement and support that made this research journey an unforgettable experience.

Finally, I would like to thank my partner for being there every step of the way…

Statement of Contributions of Collaborators

I hereby declare that this work has not been submitted or accepted for any other degree. Dr. Jules Thibault and Dr. Silvia Ochoa supervised and are co-authors of the two papers contained in this thesis.

The responsibilities of the author, in order to fulfill the requirements of this thesis, were as follows:

1. To conduct research in the area of multi-objective optimization in order to give an overview on the methods used to solve the multi-objective optimization problem. Furthermore, visualization techniques and methods to find an optimal solutions from the Pareto domain were also studied.
2. To perform multi-objective optimization on various chemical engineering applications to understand the importance of the Pareto domain.
3. To develop a model of an Acrylic Acid production plant.
4. To develop a methodology to solve multi-objective optimization using artificial neural networks as metamodels to reduce the computational load.
5. To produce two papers for publication based on the research done.
6. To write a paper-based thesis in partial fulfillment of the requirements for obtaining a Master of Applied Science in Chemical Engineering.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The process of systematically and simultaneously optimize a collection of conflicting objective functions is called Multi-Objective Optimization (MOO) [1]. This type of optimization generates a trade-off space containing a set of multiple optimal solutions, known as Pareto domain, as opposed to a unique solution as it is the case in Single-Objective Optimization (SOO). Traditionally, the multi-objective optimization problems have been solved as a single-objective problem, known as scalarization techniques [2], where the multiple objectives are combined into a single weighted sum function, for example, or using other techniques like optimizing one of the functions while the others are handled as constraints. However, these techniques rely on the assumption that the objective functions are well behaved which means that they have concave or convex shapes and are continuous [3]. The advantage of using MOO techniques instead of SOO ones is that, through the Pareto domain, it provides the decision maker (DM) with valuable information about the trade-offs existing among the different objective functions, while if the multiple objectives are combined into a single objective, only one optimal solution will be provided to the DM. However, as scalarization techniques deal with the possibility of finding the global optimum or multiple local optima, possible optimal solutions could be missed if the functions are non-convex, multi-modal or discontinuous [4]. Usually these techniques are based on deterministic algorithms and require information about the differentiability of the function which can complicate the optimization task. Another disadvantage of using scalarization techniques like the *weighted sum* is that the solution will be highly sensitive to the selection of the weights assigned to each objective.

Three goals are usually considered to solve MOO problems: (1) convergence, (2) diversity, and (3) uniformity of the solutions [5]. As only a finite number of solutions can be obtained during the optimization process, in order to maintain a reasonable computational time, a well-distributed set should be obtained so that all the possible regions of the Pareto domain are explored [5]. In fact, a good MOO algorithm should be one that will maintain diversity which will allow the decision maker to obtain solutions

that are uniformly distributed over the Pareto domain [6] yielding the maximum amount of information on the process. The Pareto domain is a very useful source of information as it does not only contain all potentially-optimal solutions to an optimization problem, but it also allows to observe all the intrinsic relationships among the different process variables that are instrumental to gain a better insight on the various aspects of the process.

As mentioned above, MOO is well suited to deal with objective functions that give rise to trade-offs between two or more conflicting alternatives. For example, consider an industrial process where the profit must be maximized by reducing the capital and operating costs. On the other hand, environmental emissions should also be minimized. However, these objectives tend to be contradictory as reducing emissions could lead to having additional equipment for the treatment of the emissions, which will affect the capital and operating costs.

Mathematically, the multi-objective optimization problem, also known as multi-criteria optimization or many-objective optimization when dealing with more than three objective functions [7], is defined in Equation (1)

$$
\begin{aligned}
&\min_{x} \; F(x) = \left[F_1(x), \ldots, F_J(x) \right], \\
&\textbf{subject to} \;\; \mathrm{h}_m(\mathrm{x}) = 0, \qquad m = 1, \ldots, M \\
&\qquad\qquad \mathrm{g}_k(\mathrm{x}) \leq 0, \qquad k = 1, \ldots, K \\
&\qquad\qquad \mathrm{x}_n^{\min} \leq \mathrm{x}_n \leq \mathrm{x}_n^{\max} \qquad n = 1, \ldots, N
\end{aligned}
\tag{1}
$$

It is desired to simultaneously minimize the vector of J objective functions $F(x)$ by finding the vector of N decision variables ($x = [x_1, x_2, \ldots, x_N]$) that leads to the best compromise among all these objective functions and that satisfy M equality and K inequality constraints, leading to the *feasible solution space* [8]. If one or more objective functions must be maximized, the negative of the said functions is minimized.

In SOO problems, the superiority of a solution with respect to other solutions is easily determined by comparing their objective function values. However, in Pareto-based optimization, the superiority of a solution will be dictated by the *dominance* of the

said objective. In fact, the Pareto domain contains all non-dominated solutions, i.e. optimal solutions, where the improvement of any one of these objectives leads to the deterioration of at least another objective [9]. A solution x_1 will dominate a solution x_2 if:

1) $\forall j.\ F_j(x_1) \leq F_j(x_2)$ where $j \in \{1, 2, \ldots, J\}$ and
2) $\exists j$ such that $F_j(x_1) < F_j(x_2)$ where $j \in \{1,2, \ldots, J\}$

Figure 1-1 presents a hypothetical solution space with three different values where objective F_1 is to be maximized while F_2 is to be minimized. Using the dominance test to determine whether a solution is dominated or not, the following is true: solutions x_2 and x_3 do not dominate solution x_1 as they are worse in F_2 but better in F_1, thus solution x_1 is a non-dominated solution with respect to the other two solutions. However, when comparing solution x_2 to solution x_3; the solution x_3 is non-dominated with respect to x_2 because x_3 is better than x_2 for both criteria.

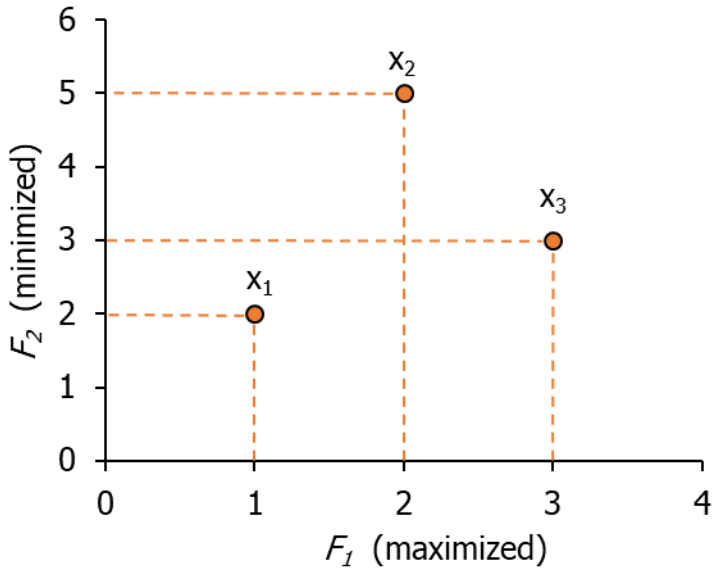

Figure 1-1. Hypothetical solution space where F_1 is to be maximized and F_2 minimized.

In order to carry out a MOO process in the field of chemical engineering, a simulation model or experimental data is required to facilitate the evaluation of the objective functions. One of the major disadvantages of explicitly solving a MOO problem is that it can be very expensive in terms of the required computational load given the

large number of model evaluations that are required to circumscribe the Pareto domain as such evaluations are typically time-consuming and expensive [10]. A common approach to tackle this problem and reduce the simulation time, is to use a surrogate model, also known as metamodel, in order to represent the underlying relationships existing between the decision process variables and the objective functions based on a first-principles model or on past evaluations. Artificial Neural Networks (ANNs) are an excellent example of metamodels that could be used to encapsulate the underlying relationships within the process data. They can model very complex nonlinear behavior, even for complex problems dealing with more than three objective functions, and once they are trained they can be used to perform a large number of simulations and determine very rapidly the Pareto domain.

The subject of this thesis is mainly focused on two aspects. The first one is to emphasize and illustrate the importance of carrying multi-objective optimization in order to circumscribe the Pareto domain. It is desired to show that the Pareto domain provides invaluable information on the process and leads naturally to a better understanding on the underlying relationships among all process variables. This is illustrated using four simple examples of processes from the field of chemical engineering. The second part of this thesis is focused on proposing an easy-to-follow methodology to solve multi-objective optimization using artificial neural networks in an effort to reduce the computational time.

1. Thesis Objectives

The main objectives of this thesis are:

1) To give a quick overview of the different methods available to solve multi-objective optimization problems, the multiple visualization techniques and the several methods for selecting an optimal solution.

2) To understand the importance of multi-objective optimization and the Pareto domain through several chemical engineering applications, more specifically the

design of a PI controller, an SO_2 to SO_3 reactor, a distillation column and finally of an acrolein reactor.

3) To propose an easy-to-follow methodology to solve multi-objective optimization based on first-principles model of an acrylic acid production plant in order to train an artificial neural network that is used as a metamodel to reduce the computational load of the optimization problem.

2. Structure of the Thesis

This thesis consists of four chapters. The first chapter presents a general introduction on the MOO field and the different research objectives that were considered during the course of this research. Chapters 2 and 3 contain papers that will be submitted to scientific journals based on the two main topics of the thesis. Due to the paper-based structure, it is possible to find that some concepts will be repeated through this work. As for Chapter 4, it contains a general discussion tackling the main conclusions and recommendations based on the work that was done. Finally, a more detailed description of the acrylic acid production plant model is given in Annexes A and B.

References

[1] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidiscip. Optim.*, vol. 26, no. 6, pp. 369–395, 2004.

[2] A. Nikjoofar and M. Zarghami, "Water Distribution Networks Designing by the Multiobjective Genetic Algorithm and Game Theory," *Metaheuristics Water, Geotech. Transp. Eng.*, pp. 99–119, 2013.

[3] H. Halsall-Whitney and J. Thibault, "Multi-objective optimization for chemical processes and controller design: Approximating and classifying the Pareto domain," *Comput. Chem. Eng.*, vol. 30, no. 6–7, pp. 1155–1168, 2006.

[4] E. Thomas F., D. M. Himmelblau, and L. Lasdon, *Optimization of Chemical Processes McGraw-Hill Chemical Engineering Series*, 2nd ed. New York: McGraw-Hill, 2001.

[5] M. Elarbi, S. Bechikh, L. Ben Said, and R. Datta, “Multi-objective Optimization: Classical and Evolutionary Approaches,” in *Recent Advances in Evolutionary Multi-objective Optimization*, Springer, Cham, 2017, pp. 1–30.

[6] A. Konak, D. W. Coit, and A. E. Smith, “Multi-Objective Optimization Using Genetic Algorithms: A Tutorial,” *Reliab. Eng. Syst. Saf.*, vol. 91, pp. 992–1007, 2006.

[7] B. Li, J. Li, K. E. Tang, X. Yao, and K. Tang, “Many-objective evolutionary algorithms: A survey,” *ACM Comput. Surv*, vol. 48, no. 13, pp. 1–34, 2015.

[8] T. Erfani and S. V. Utyuzhnikov, “Directed search domain: A method for even generation of the Pareto frontier in multiobjective optimization,” *Eng. Optim.*, vol. 43, no. 5, pp. 467–484, 2011.

[9] C. A. Coello Coello, “Evolutionary multi-objective optimization: a historical view of the field,” *IEEE Comput. Intell. Mag.*, vol. 1, no. 1, pp. 28–36, 2006.

[10] M. T. M. Emmerich and A. H. Deutz, “A tutorial on multiobjective optimization: fundamentals and evolutionary methods,” *Nat. Comput.*, vol. 17, no. 3, pp. 585–609, 2018.

Chapter 2. Pareto Domain: An Invaluable Source of Process Information

Geraldine Cáceres Sepúlveda[1], Silvia Ochoa[2], and Jules Thibault[1*]

[1]Department of Chemical and Biological Engineering, University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
[2] SIDCOP Research Group-Departamento de Ingeniería Química
Universidad de Antioquia, Medellín, Colombia

Abstract

Due to the highly competitive market and increasingly stringent environmental regulations, it is paramount to operate chemical processes at their optimal point. In a typical process, there are usually many process variables (decision variables) that need to be selected in order to achieve a set optimal objectives for which the process will be considered to operate optimally. Because some of the objectives are often contradictory, Multi-Objective Optimization (MOO) can be used to find a suitable trade-off among all objectives that will satisfy the decision maker. The first step is to circumscribe a well-defined Pareto domain, corresponding to the portion of the solution domain comprised of a large number of non-dominated solutions. The second step is to rank all Pareto-optimal solutions based on some preferences of an expert of the process, this step being performed using visualization tools and/or a ranking algorithm. The last step is to implement the best solution to operate the process optimally. In this paper, after reviewing the main methods to solve MOO problems and to select the best Pareto-optimal solution, four simple MOO problems will be solved to clearly demonstrate the wealth of information on a given process that can be obtained from the MOO instead of a single aggregate objective function. The four optimization case studies are the design of a PI controller, an SO_2 to SO_3 reactor, a distillation column and an acrolein reactor. Results of these optimization case studies show the benefit of generating and using the Pareto domain to gain a deeper understanding of the underlying relationships between the various process variables and performance objectives.

Nomenclature

C	Ratio of R over R_{min}	-
CAPEX	Capital cost	M$
C_{BM}	Bare module cost of the equipment	$
C_p	Purchased cost of the equipment	$
C_{sb}	Souders Brown parameter	-
D_C	Diameter of the column	m
DM	Decision maker	-
DPEA	Dual Population Evolutionary Algorithm	-
E_o	Overall tray efficiency	%
f	Fraction of the flooding velocity	-
f(∞)	Final output of the system	-
f(0)	Initial output of the system	-
f(t)	System response	-
F_F	Foaming parameter	-
F_{HA}	Hole-area factor	-
F_M	Material factor depending on the material of construction	-
F_P	Pressure factor	-
F_q	Quantity factor based on N	-
FOPDT	First Order Plus Dead Time model	-
F_{ST}	Surface tension parameter	-
H_C	Height of the column	m
H_P	Plate spacing	m
ISDU	Integral of the squares of the differences in the manipulated variable	-
ITAE	Integral of the time weighted Absolute Error	-
K_C	Controller gain	-
K_P	FOPDT process gain	-
N	Number of trays	-
NFM	Net Flow Method	-
N_{min}	Minimum number of trays	-
OPEX	Operating cost	M$/year
OZ		s

	Integral of the time the output variable spent outside the ±5%	
P	envelope	$kmol/m^3h$
PI	Productivity of the reactor	-
PID	Proportional-integral controller	-
q	Proportional-integral-derivative controller	-
Q	Distillation feed thermal quality	kJ/h
R	Rate of heat transfer	-
R_{min}	Ratio of R over R_{min}	-
t	Minimum reflux ratio	s
u	Time	-
U	Manipulated variable	W/m^2K
W	Overall heat transfer coefficient	kg
X	Catalyst weight	%
x_i	Conversion of the reactor	-
	Molar fraction of component i	

Greek symbols

α	Relative volatility	-
Δu_t	Change of the manipulated variable at time t, i.e. (u_t-u_{t-1})	-
ε_t	Instantaneous error	-
θ	Underwood constant	-
θ_P	FOPDT process dead time	s
μ	Viscosity of the fluid in the column	cP
ρ	Density	kg/m^3
τ_I	Integral time of the controller	s
τ_P	FOPDT process time constant	s

1. Introduction

In our everyday life, a large number of decisions being made can be considered as multi-objective optimization problems because in most cases our decisions represent some types of trade-offs between two or more alternatives. These alternatives, also known as objective functions, are very often contradictory [1]. The same is true for the operation of industrial processes where multi-objective optimization is omnipresent. Multi-objective optimization, also known as multi-criteria optimization, deals with problems that are represented mathematically by Equation (1):

$$\begin{aligned} &\min_{x} \; F(x) = \left[F_1(x), ..., F_J(x) \right], \\ &\textbf{subject to} \;\; \mathrm{h}_m(\mathrm{x}) = 0, \qquad m = 1, \ldots, M \\ &\qquad\qquad \mathrm{g}_k(\mathrm{x}) \le 0, \qquad k = 1, \ldots, K \\ &\qquad\qquad \mathrm{x}_n^{\min} \le \mathrm{x}_n \le \mathrm{x}_n^{\max} \qquad n = 1, \ldots, N \end{aligned} \tag{1}$$

Equation (1) states that it is desired to simultaneously minimize the vector of J objective functions $F(x)$ by finding the vector of N decision variables ($x = [x_1, x_2, \ldots, x_N]$) that leads to the best compromise among all these objective functions. This optimal solution must satisfy a series of M equality and K inequality constraints. At the same time, the values of the N decision variables that satisfy this problem are bounded within the *feasible solution space* [2]. If one or more objective functions need to be maximized, then the negative of that function is minimized. As mentioned above, solutions to a multi-objective optimization problem are nearly invariably trade-offs between the various objectives since an improvement of one objective may be detrimental to another objective [3]. These solutions are known as non-dominated solutions and are said to be *Pareto-optimal* so that the conglomerate of these solutions gives rise to the *Pareto domain* or *Pareto optimal set* [4]. All other solutions not located in the Pareto domain are considered dominated, suboptimal, or locally optimal [3].

The Pareto domain, that is the set of non-dominated solutions, is ideally represented by a large number of well-distributed solutions covering the entire solution space. With respect to the concept of dominance [5], mathematically a solution x_1 is said to dominate another solution x_2 if:

1) $\forall j.\ F_j(x_1) \leq F_j(x_2)$ where $j \in \{1, 2, \ldots, J\}$ and

2) $\exists j$ such that $F_j(x_1) < F_j(x_2)$ where $j \in \{1,2, \ldots, J\}$

The solution to a multi-objective optimization problem boils down to determining the set of Pareto-optimal solutions. Over time, different methods and approaches have been developed in order to solve MOO problems. Section 2 will provide an overview of the most widely used methods.

The term *many-objective optimization* [6] is used by many for optimization problems with more than three objective functions, even though the term multi-objective optimization is often used. This distinction was made to draw attention to the challenges that are faced when dealing with more than three objective functions [7]. According to Emmerich and Deutz, reasons for this distinction are that 1) problems with many objectives have a Pareto front which cannot be visualized in conventional 2D or 3D plots so that other approaches to deal with this limitation are needed; 2) the computation time for many indicators and selection schemes become computationally more difficult; and 3) the ratio of non-dominated points tends to increase rapidly with the number of objectives [7].

To select the optimal region of operation, once the Pareto domain is approximated, the decision maker (DM) needs to explore all the available solutions and based on his knowledge of the process, makes his final selection. However, the Pareto domain can contain a very large number of candidate solutions, and an exhaustive search is difficult and inefficient when the DM seeks to identify and negotiate trade-offs between some of the conflicting objectives [8]. Usually, the Pareto domain surface of a many-objective optimization problem is commonly projected in a 2D space which renders the task of analyzing the results more complex. To account for these issues, multidimensional data visualization techniques are often used to reveal the different trade-offs among the objectives in the solution space as they provide the DM with a meaningful method to analyze the Pareto set and select appropriate solutions [9]. Additionally, there are methods available for selecting an optimal solution according to the preferences of the DM of the process (see Section 3 of this chapter for more information on these methods).

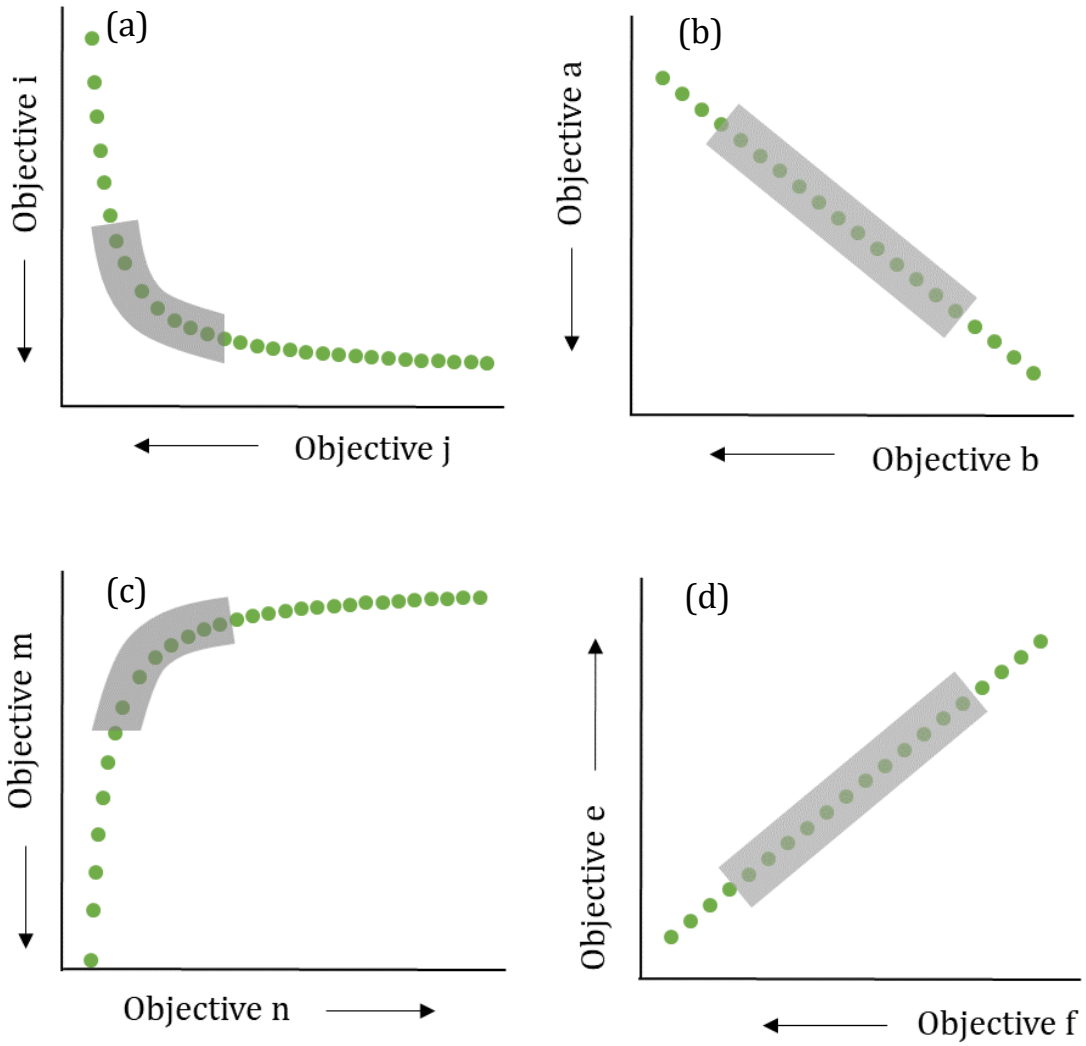

Figure 2-1. Simple examples of Pareto domain and trade-offs for different objective pairs.

By distinguishing the shape of the Pareto domain, the DM can prioritize the exploration of the search space as the shape of the curve dictates the extent of the trade-off between objective pairs [10]. Figure 2-1 displays hypothetical Pareto domains obtained for different objective pairs. In each case, the preferred direction for each objective is indicated by the arrows and the region that will usually provide the best trade-offs between the objectives is highlighted in gray. The shape of the curve will have a direct impact on the DM selection. For instance, for concave or convex plots shown in Figures 2-1(a) and 2-1(c), respectively, the gray shaded region will be of interest since all the solutions in this region result in improvement in one of the objectives for the least deterioration in the other objectives in comparison to the solutions outside of the shaded

region. In comparison, the DM will have to explore a larger region of the Pareto front in Figures 2-1(b) and 2-1(d) as changes are nearly identical in both objectives for every solution [10]. However, it is important to recall that all the solutions in the Pareto domain are still optimal and they were obtained without any bias from the DM. Section 4 of this paper is centered on the importance of the Pareto domain as it does not only contain all potentially-optimal solutions to the optimization problem but also allows to observe all the intrinsic relationships among the different process variables that are instrumental to gain a better insight on the various aspects of the process. This will be achieved through four different chemical engineering applications.

2. Methods to Solve Multi-Objective Optimization

Solving a multi-objective optimization problem signifies to find the Pareto domain containing all the potentially-optimal solutions to the problem. All solutions located outside the Pareto domain are dominated meaning that at least one solution can be found within the Pareto domain that will be better for all objectives. In general, optimization methods can be grouped into two major categories: classical methods that usually use deterministic approaches, and evolutionary algorithms based on stochastic algorithms, where a set of candidate solutions gradually evolve to the Pareto set [11]. Methods of the former class will often reduce a MOO problem into an aggregate objective function, such as a weighted sum of all objectives. The latter class, on the other hand, will consider all objective functions with the goal to optimize them simultaneously. Evolutionary methods start by generating a random or uniformly-distributed initial population. Through an iterative procedure via an analysis of the degree of dominance of all current solutions, where some members of the population are allowed to mate to generate the next population. To ensure the circumscription of a well-balanced and diversified population, some operators such as the selection, mutation and crossover operators are used. This evolutionary procedure is continued until some stopping criteria are satisfied. The stopping criteria could be reaching a given number of iterations, obtaining a desired number of non-dominated solutions, reaching a time limit, and satisfying some constraint tolerance or fitness values [11].

A good algorithm to solve MOO problems should be one that will yield solutions with a good diversity in order to avoid having solutions mostly located in a very restricted region of the feasible solution space. This task is as important as the task of converging to the true Pareto-optimal set [12]. Diversity will allow the decision maker (DM) to obtain solutions that are uniformly distributed over the Pareto domain [13] which provides the maximum amount of information [14].

Usually, once the Pareto domain is approximated by a large number of non-dominated points, the set will be presented to the DM as a whole so that the selection of the Pareto-optimal objective vector and its corresponding decision vector [15] can be made. More information on how to select an optimal solution and how to visualize the Pareto domain will be given in subsequent sections.

2.1 Deterministic Algorithms

2.1.1 Scalarization Techniques

There are plenty of methods that have been proposed to solve MOO problems. One of the most used methods to calculate the Pareto front are to transform a multi-objective optimization problem into a single-objective optimization problem. The *weighted sum* (WS) approach attempts to minimize the sum of the positively normalized, weighted, single-objective scores [16]. The well-known drawbacks of this method are that 1) the solution is strongly correlated to the selection of the weights, 2) the optimal solution distribution is often not uniform, and 3) optimal solution in non-convex regions are not detected. [4], [17] The *adaptive weighted sum* (AWS) was proposed as an attempt to counteract these limitations. It uses similar mathematical formulation to the WS method. The most important feature of this method is that it adaptively focuses on unexplored regions of the Pareto set by specifying additional constraints for the regions where further refinement is required [18]. Another classical method consists in optimizing one chosen objective function and transforming the others into constraints.

Target vector optimization is another possible method. It requires a vector of goal values so that the optimization is driven towards the shortest distance between any

candidate solution and the goal vector [16] [19]. *Goal programming* is also a candidate method. In this case, one objective is minimized while the remaining objectives are constrained to be less than the target values [20]. One major drawback of this method is the selection of appropriate goals for the constraints [16]. It is important to bear in mind that solving a multi-objective optimization problem by means of single objective optimization techniques will always result in a single solution.

2.1.2 Pareto Based Deterministic Multi-Objective Techniques

Pareto-based deterministic methods have been proposed to solve MOO problems. The *Normal Boundary Intersection* (NBI) method was proposed by Das and Dennis [21]. The NBI method produces well-distributed solutions, and it is easily scalable to n-dimensional problems. It generates both non-Pareto and locally Pareto solutions, so that it requires a filtering procedure [17]. In addition, the NBI method might be non-robust since the feasible domain is reduced to a line [2] [17] since the addition of equality constraints forces the solution to lie on a line normal to the Utopia line [18].

Another method that falls under this category is the *Normal Constraint* (NC) method proposed by Messac et al. [22]. It is only based on inequality constraints and it is one of the easiest and fastest methods to implement and understand [18]. Akin to the NBI method, it has the inconvenient of finding dominated solutions, so that a filtering step must also be added. One major drawback of both methods is that they can encounter significant difficulties in the case of a disconnected frontier [2].

The *Physical Programming* (PP) method introduced by Messac [23] and Mattson [24] is capable of generating an evenly distributed Pareto set. The decision maker will express his preferences with respect to each objective using four different classes [24]. However, the optimal selection of these parameters will rely on preliminary information on the location of the Pareto frontier [2]. A more comprehensive survey of the various continuous deterministic methods are presented by Marler and Arora [20].

2.2 Stochastic Algorithms

Stochastic optimization methods are sometimes based on concepts that we can encounter in everyday life [25] and they are usually divided into evolutionary and non-evolutionary methods. These methods are easy to implement since no derivative information is required to solve the optimization problem [26]. This type of methods implement random decision when looking for the optimum. However, the convergence to the optimum could be slow due to the working principles of the methods.

Only evolutionary algorithms will be considered under this category in this paper, more specifically, multi-objective evolutionary algorithms (MOEAs). When solving the MOO problem, these algorithms consider all the objective functions simultaneously to generate the Pareto domain. Evolutionary Algorithms (EAs) are population-based, bio-inspired, black-box optimization methods that do not need assumptions of continuity or differentiability. These methods are based on Darwin's theory of survival of the fittest [27] so that they use principles from natural evolution, such as selection, recombination, and mutation to evolve a population of individuals, i.e. decision variable vector, towards optimal or near-optimal solutions [28] that constitute the Pareto domain. MOEAs were first developed in the 90s, but this topic gained strength in early 2001 when Deb published a book solely on this topic [29] and numerous methods have been developed since then [7].

Evolutionary algorithms have a number of clear advantages over other classical methods. They are not sensitive to non-smoothness of the objective functions, are efficient in finding a global optimum [2] and they have an inherent robustness to noise [30]. For instance, it is suggested that multi-objective optimization might be a problem area where EAs do better than other blind search strategies [31]. However, as other methods, EAs have their own drawbacks. A major one is that in order to generate an even set of optimal solutions, a huge number of solutions are considered which could lead to redundancy. This also means that the process to solve MOO problems is time-consuming [2]. It is also well known that these stochastic and population-based optimization algorithms tend to produce some variants of the best solution, and the difference in

properties and quality of these solutions will depend on the algorithm employed, the length of the run and how diversity is maintained through the process [32].

One of the most well-known and widely-used evolutionary method is the *Genetic Algorithms* (GAs). They were initially developed by Holland in the 1970s [33] and the idea behind these algorithms is to mimic the natural selection and survival of the fittest. Some genetic operators like crossover, mutation and selection are used in order to represent these natural processes. GAs are able to search in a large space to make candidate solution sets to the optimization problem evolve to finally keep a set of good solutions in the last generations [34] [35].

There are other two popular methods that are similar to GA: *Particle Swarm Optimization* (PSO) and *Differential Evolution* (DE). Although GAs are more well-established due to their earlier development, these two methods have started to attract more attention during the past years, especially for continuous optimization problems [36]. PSO was first introduced by Kennedy and Eberhart in 1995 [37] and it was based on the emulation of the collective behavior of bird flocks [38]. As for the DE method, it was introduced that same year by Price and Storn [39]. This method was named after the differential operator that they proposed to create new offspring from parent chromosomes in an attempt to replace the crossover and mutation operators of the GA [38]. Voratas performed a good comparison of the performance of the three algorithms based on the similarities and differences among the three methods [36].

3. Pareto Domain

3.1 Methods for Selecting an Optimal Solution

Multi-objective optimization based on Pareto optimality usually consists of two main steps: the determination of the Pareto-optimal set and the presentation of this set to an expert or decision maker for the selection of the optimal solution. It is important to emphasize that the Pareto domain is circumscribed without any bias, since it is determined without assigning any preferences to the objectives, apart from specifying if the said objectives are to be minimized or maximized. It is evident that some of the

Pareto-optimal solutions are better that other solutions according to the DM preferences or the process requirements. In this sense, a method that allows to rank all Pareto-optimal solutions according to those preferences is required.

As stated earlier, the MOO process will generate many optimal solutions that will be equally good when only considering the given objectives. The Pareto domain containing these solutions will provide a deeper insight into the trade-offs that will inevitably exist between the different objectives such that a single optimal solution can be selected. A very exhaustive study by Wang and Rangaiah was done on the different methods for selecting an optimal solution to a MOO problem from the Pareto-optimal set according to the preferences that a DM has on the process [40]. These methods all fall under the area of multi-criteria decision making (MCDM), which is a well-known topic of decision making [41].

One of the most common methods is the *Technique for Order of Preference by Similarity to Ideal Solution* (TOPSIS) method. It was first developed in 1981 by Yoon and Hwang [42]. Its working principle is that the chosen alternative should have the shortest distance from the ideal solution and the farthest from the negative-ideal solution [43]. The ideal solution is a combination of the best value of each objective among the given optimal solutions whereas the negative-ideal solution is a combination of the worst value of each objective [40]. Another method of interest is the *Gray Relational Analysis* (GRA) first developed by Deng [44]. In this method, all the alternative's performances are compared to an ideal network obtained by selecting the best value of each objective through the grey relational coefficient (GRC). The one with the highest value will be the best choice. [45] One of the advantages of this method is that it does not require weights or any other input from the user [40].

The *Simple Additive Weighting* (SAW) basic concept consists on calculating an evaluation score for each alternative by multiplying the scaled value of the solution with the weight assigned by the decision maker and followed by the summation of the products for all criteria [46] [47]. The solution with the highest evaluation score will be the one selected. Additionally, one can use the *Net Flow Method* (NFM), which is an

amalgam of the *Elimination and Choice Translating Priority III* (ELECTRE III) and the *Preference Ranking Organization Method for Enrichment of Evaluations* (PROMETHEE) methods [48]. It consists of ranking all Pareto-optimal solutions based on the difference of the extent of one objective outranking other objectives and vice-versa. The largest ranking score will dictate the best solution. This method requires the user to define four parameters for each objective criterion, namely the relative weight (W_j), the indifference threshold (Q_j), the preference threshold (P_j), and the veto threshold (V_j) [48] [49].

Other methods are also available to select an optimal solution, like the *Linear Programming Technique for Multidimensional Analysis of Preference* (LINMAP) [50] [51], the *Viekriterijumsko Kompromisno Rangiranje* (VIKOR) method [52], the *Rough Set Method* (RSM)[48] [53], *Faire Un Choix Adéquat* (FUCA) [54], *Elimination and Choice Translating Priority II* and III (ELECTRE II and III) methods [55].

3.2 Visualization Techniques to Evaluate the Pareto Domain

As previously discussed, the Pareto domain consists of all non-dominated solutions to a specific optimization problem and trade-offs between these objectives will always exist since improvement of one objective can result in the deterioration of another one. There are many measures available to assess the quality of an approximation set or to determine the optimal solution according to the DM preferences, as mentioned earlier. However, no measure is as effective as the visualization of the Pareto in order to 1) estimate the location, range, and shape of the Pareto domain, 2) assess the different trade-offs between objectives, 3) select preferred solutions, etc. [56] Visualization tools are valuable and provide DM with a meaningful method to analyze the Pareto set to determine the intrinsic relationships among the conflicting variables and select good trade-off solutions [9] [10].

A Pareto domain that is obtained employing MOEAs can often be very large. Thereby, multidimensional data visualization methods that are able to handle large data sets in terms of visualization capability as well as computational complexity are needed [56]. These methods should be scalable to multiple dimensions, as well as easy to use and

understand [56]. Nevertheless, as the number of objectives increases in a MOO problem, the visualization task becomes increasingly difficult when it comes to revealing all trade-offs that exist in the solution space [10]. Visualization of a Pareto domain is indeed considered a difficult task when the problem at hand spans more than three objective functions [57]. In fact, the Pareto front of up to three objectives can be easily visualized using *scatter plot matrices* [58], for example. When it comes to four or five objectives, other techniques like *decision maps*, *polyhedral approximation*, or other visualization techniques can be used [15]. Some methods, known as interactive visualization techniques, may be able to handle the display of more than three objectives, nevertheless, the cognitive effort required for this task will be higher [15] [32].

Two of the most commonly used methods for graphical representation of the Pareto domain are *parallel coordinate plots* [56] [59] [60] and *scatter plot matrices* [60]–[62]. Both methods are scalable to represent any number of objectives and data size since they present the objectives by pairs in each graph. However, it results in a notable cognitive effort in order to fully explore the trade-off space and reveal all the trade-offs among the objectives [10].

Other advanced multi-dimensional data visualization methods have been proposed to represent Pareto domains with more than four objectives in a more efficient way. These techniques include *chord diagrams* [63] [64], *self-organizing maps (SOM)* [57] [30], *level diagrams* [9] [65], *heat maps* [66] [67], i*nteractive decision maps* [68], *hyper-radial visualization* [69] , and many more.

Filipič and Tušar present a comprehensive review of the existing visualization methods used in MOO where the methods are divided into two groups: 1) general multidimensional data visualization methods and 2) methods specifically designed for visualization of approximation sets [56].

3.3 Interactive Scatter Plot Matrices

Each visualization technique has their own advantages and disadvantages. The main advantages of using scatter plot matrix is that it is a simple, fast and robust method,

and it allows to visualize a large number of data points while being computationally inexpensive [56]. One major drawback of this method is that its use can incur in a loss of information due to the dimensionality reduction to 2D or 3D figures [58].

To alleviate this shortcoming, a program was developed in order to help with trade-off space exploration and understand the different trade-offs that could exist among the variables. This program was developed in MATLAB by Paul Geiger [70] and it consists of interactive scatter plot matrices. Its intended purpose is to plot the Pareto domain of multiple objectives in 2D or 3D plots to allow the exploration of the trade-off space by enabling the user to select a point or a region in one of the 2D/3D plots. The corresponding points in that plot and all the other plots will then be highlighted. This feature allows the user to understand the different trade-offs between the variables and more easily observe the intrinsic relationships among all of them.

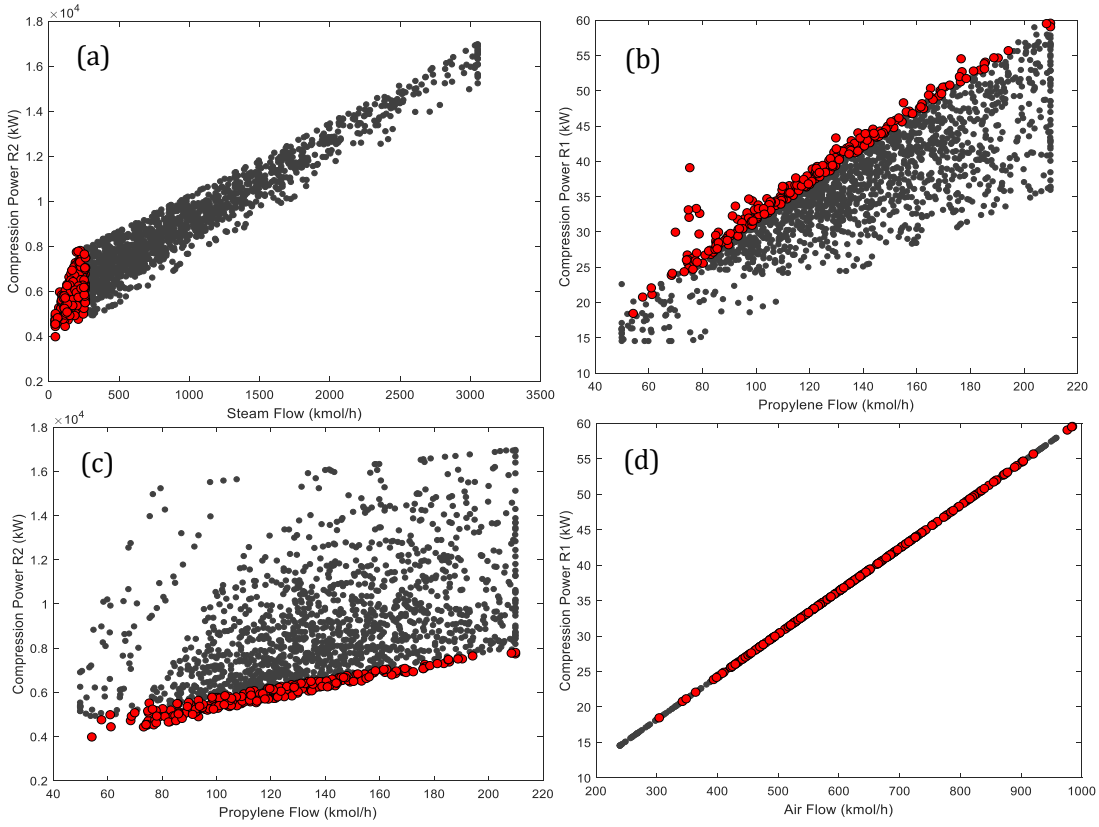

Figure 2-2. Illustration of scatter plot matrices of the Pareto domain for the MOO of an acrylic acid production plant involving eight objective functions.

This program is very user-friendly. The code, as well as the manual of instructions, are available in the additional material associated to this thesis. When scatter plot matrices are used to represent the Pareto domain obtained for more than three objective functions, the multiple dimensions are projected in 2D or 3D spaces. It also allows visualizing at the same time the corresponding decision variables. Figure 2-2 shows the Pareto domain of a MOO problem solved for an acrylic acid production plant. The optimizations for this case is performed to simultaneously optimize eight objective functions. The highlighted red zone of Figure 2-2(a-d) represents the same Pareto domain region in all the graphs and are a sample on how this program works. In cases where more than 3 objectives must be optimized, having such a tool will prove to be very helpful in analyzing the Pareto domain.

4. Importance of the Pareto Domain: Case Studies

The main goal behind performing multi-objective optimization is to find the Pareto domain, which contains all the potentially optimal solutions to a specific problem. However, it has many other interesting uses. For instance, it can be used to help the decision maker in having a better understanding of the process and the relationships that may exist among the process variables. In this section, the importance of the Pareto domain is analyzed considering four different chemical engineering applications. The first case study optimization is carried on a proportional-integral (PI) controller in order to find optimal controller parameters that will minimize three performance criteria: ITAE, ISDU and OZ. The next case study will be done on an SO_2-SO_3 reactor system that is subject to a series of constraints. The main goal is to maximize the productivity and conversion of the reactor while minimizing the catalyst weight required. Subsequently, a distillation column is solved analytically to be able to analyze well-known information under a new perspective. Finally, a fourth case study is performed on a first reactor of a system of two reactors in series, responsible for the production of acrylic acid. As it is the case for the SO_2 to SO_3 reactor system, productivity and conversion will be maximized for this reactor, and a constraint on the oxygen concentration handled as a soft constraint, is

selected as the third criterion to minimize. All the decision variables and objectives for the four case studies are defined in Table 2-4 in Appendix B of this chapter.

The Dual Population Evolutionary Algorithm (DPEA) is used [70] to circumscribe the Pareto domain for each case study and to further classify the solution space, the Net Flow Method (NFM) is used [48]. All the parameters that were used for this method are presented in Table 2-5 in Appendix C. Based on the NFM ranking, the Pareto domain of each case study is divided into four different regions: i) the best solution in red, ii) Pareto-optimal solutions ranked in the top 5%, iii) solutions in the next 45% slice and iv) the remaining 50% of the solutions. The Net Flow Method is relatively robust and changes in the weights of the objective will not incur in major changes of the optimal zones [48].

4.1 PI Controller

Proportional-integral-derivative (PID) control is undoubtedly the most-used feedback control algorithm. In fact, because of the presence of noise, the derivative action will often not be implemented such that the controller is reduced to its proportional-integral (PI) form [72]. This type of controller has a logical foundation, is simple to understand and implement, and provides very good control for a wide range of industrial applications [73]. Once the proper form of the controller has been selected, the next challenge is to properly tune the controller to operate optimally despite potential disturbances and set point changes.

The first case study to highlight the insight that the Pareto domain can provide is focussed on the tuning of a PI controller for the control of a first-order-plus-dead-time process. The form of the PI controller used in this investigation is defined in Equation (2) in the Laplace domain.

$$G_C(s) = \frac{U(s)}{\varepsilon(s)} = K_C\left(1 + \frac{1}{\tau_I s}\right) \qquad (2)$$

A linear First Order Plus Dead Time (FOPDT) process model (Equation (3)) is very frequently used in the process industry to adequately approximate a higher-order system, which is then used to tune the controller. For the current case study, the process

gain, time constant and dead time of the FOPDT system were K_P = 1.5, τ_P = 5 and θ_P = 1, respectively.

$$G_P(s) = \frac{K_P e^{-\theta_P s}}{\tau_P s + 1} \quad (3)$$

To tune the PI controller for a change in set point, a multi-objective optimization is performed. The multi-objective optimization problem thereby consists of two decision variables, namely the controller gain K_C and the integration time τ_I, and three objective functions to be minimized simultaneously. The first objective function used to assess the controller performance is the Integral of the Time weighted Absolute Error (ITAE), which penalizes more heavily errors that occur late in time. The second criterion is the integral of the time the process output y(t) spent outside the ±5% zone around the set point (OZ), which favours a fast dynamic response. The third objective function is the Integral of the squares of the differences in the manipulated variable (ISDU), which ensures that the variation of the manipulated variable is smooth. The three objective functions, defined in Equations (4)-(7), all need to be minimized as mentioned above.

$$ITAE = \int_0^t t\,|\varepsilon_t|\,dt \quad (4)$$

$$OZ = \int_0^t g(f(t))dt \quad (5)$$

$$\text{where } g(f(t)) = \begin{cases} 1 & |f(t) - f(\infty)| \geq 0.05|f(0) - f(\infty)| \\ 0 & |f(t) - f(\infty)| < 0.05|f(0) - f(\infty)| \end{cases} \quad (6)$$

$$ISDU = \int_0^t \Delta u_t^2 dt \quad (7)$$

The Pareto domain for the PI controller was circumscribed for a population of 2000 non-dominated solutions and the Net Flow Method was used to rank these 2000 Pareto-optimal solutions. Figure 2-3 presents the ranked Pareto domain for the two decision variables and the three objective functions. The plot of K_C as a function of K_C/τ_I

is presented in Figure 2-3(a), which has been graphed in a way to decouple the proportional and integration actions and better see the relationship between the two controller actions. One can readily observe the very strong linear correlation between the two decisions variables. This plot indicates that by choosing a controller gain K_C from the Pareto domain, only a very narrow range of the integration action K_C/τ_I will lead to optimal performance of the controller. The reverse is also true. The linear relationship between the two controller actions suggests that the PI controller could in fact be reduced to a single parameter as the other controller parameter needs to satisfy this linear relationship for the process to remain optimal [74]. The final selection of the controller parameters remain in the hands of the decision maker who has a better understanding of the intrinsic dynamics of the process. Ranking and visualization techniques are additional techniques that the decision maker can used to assess the Pareto domain. In this case study, the NFM was used and the best controller parameters in order to simultaneously minimize the ITAE, OZ and ISDU are K_C = 0.90 and τ_I = 3.8.

Table 2-1. Controller tuning parameters and respective objective criteria.

Method	Criteria	K_C	τ_I
Ziegler and Nichols, ZN [75]	1/4 decay	$\frac{0.9}{K}\left(\frac{\tau}{\theta}\right)$	3.33θ
Cohen and Coon, CC [76]	1/4 decay	$\frac{1}{K}\left(0.9\frac{\tau}{\theta}+\frac{1}{12}\right)$	$\theta\left(\frac{30+3(\theta/\tau))}{9+20(\theta/\tau)}\right)$
Smith and Murril, SM [77]	ITAE	$\frac{0.586(\theta/\tau)^{-0.916}}{K}$	$\frac{\tau}{1.03-0.165(\theta/\tau)}$
Skogestad, SK [78]	IMC (if $\tau \leq 8\theta$)	$\frac{0.5}{K}\frac{\tau}{\theta}$	τ
Hägglund and Åström, HA [79]	Robustness	$\frac{0.14}{K}+\frac{0.28\tau}{\theta K}$	$0.33\theta+\frac{6.8\theta\tau}{10\theta+\tau}$
Chien, Hrone and Rewick, CH [80]	No overshoot	$\frac{0.35}{K}\frac{\tau}{\theta}$	1.2τ
Smith and Corripio, SC [81]	5% overshoot	$\frac{\tau}{2K\theta}$	τ
Tavakoli et al., TK [82]	MOO	$\frac{5}{11}\frac{\tau}{K\theta}$	$\frac{35}{6}\theta$

The plots of the three objective functions for the ranked Pareto domain are presented in Figures 2-3(b)-(c) in two-dimensional projections. In Figure 2-3(b), the existence of a compromise between the ITAE and ISDU criteria is truly highlighted, where a decrease in ITAE is made at the expense of an increase in ISDU and vice versa. On the other hand, ITAE and OZ criteria in Figure 2-3(c) are positively correlated as they both aim to achieve a fast response time. The Pareto domain derived for the feedback control of the FOPDT process contains only potentially-optimal solutions, i.e. only non-dominated solutions. The solutions obtained with eight common tuning algorithms based on a FOPDT process were also determined and plotted alongside the Pareto domain to compare them with the results obtained with the MOO. The equations of the eight different algorithms are given in Table 2-1 and the corresponding solutions are plotted in Figure 2-3.

The PI controller parameters for four different methods: Smith-Murril, Skogestad, Smith-Corripio and Tavakoli, lie on or in the immediate vicinity of the Pareto front. However, the parameters determined with the Ziegler-Nichols and Cohen-Coon algorithms are the ones that lie the farthest from the Pareto domain thus yielding poorer performance compared to the other methods mostly for the ISDU criterion. The objective function of the latter two algorithms is entirely different than the objective functions of this investigation. Indeed, the objective function of the Ziegler-Nichols and Cohen-Coon algorithms was derived to obtain a process response with a one-quarter decay ratio, which would obviously lead to more oscillatory responses and higher ISDU [83]. On the other hand, all algorithms gave solutions that are located on or very close to the Pareto domain projected in the ITAE-OZ space (Figure 2-3(c)). It is important to mention that in this investigation the Pareto domain was obtained for a specific FOPDT process whereas the algorithms of Table 2-1 were derived based on a large number of first-order processes over a wide range of process gain, time constant and dead time. Some discrepancies for some algorithms are therefore expected.

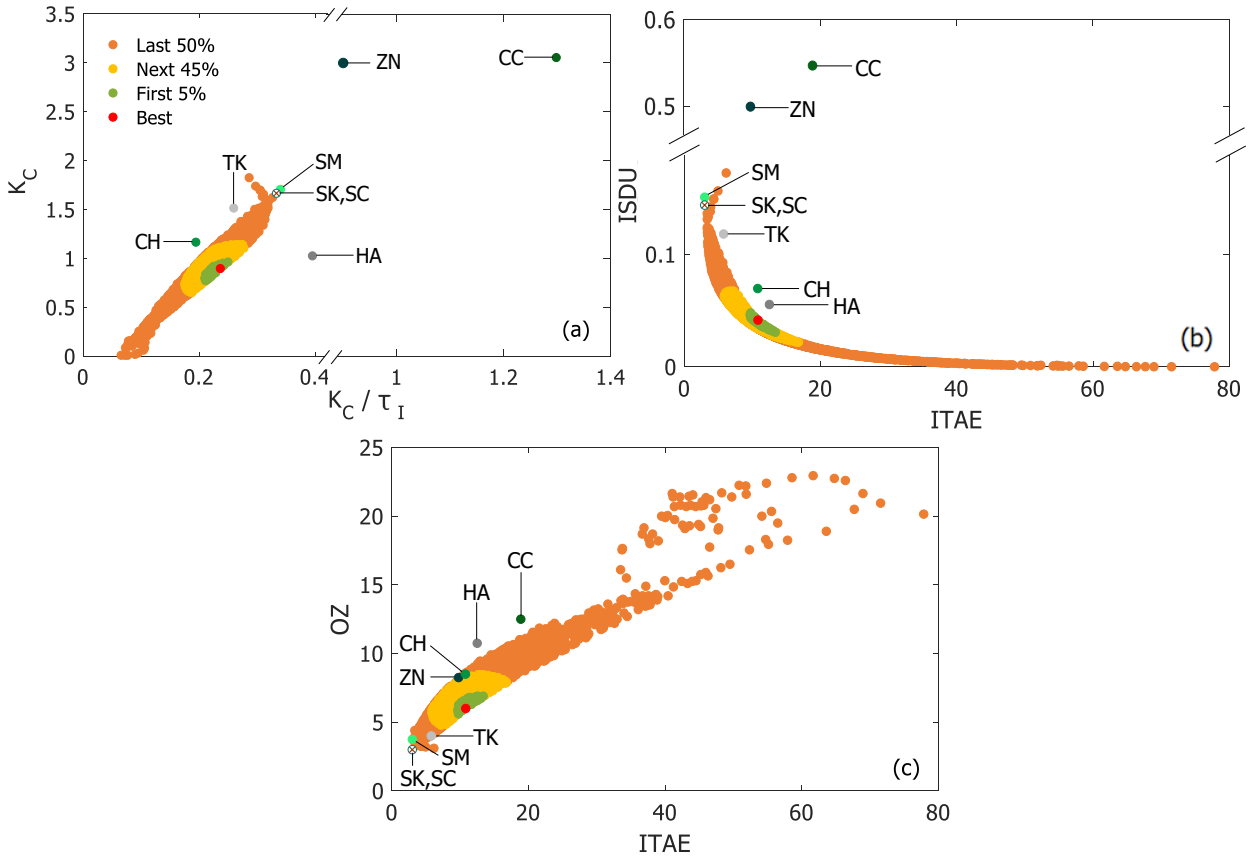

Figure 2-3. Pareto domain of the tuning of a PI Controller for a specific FOPDT process ranked with Net Flow along with the solutions obtained with different simple PI controller tuning algorithms.

4.2 SO_2 to SO_3 Reactor

Sulphuric acid is the manufactured chemical product with the highest rate of production in the world due to its myriad of applications. For the production of sulphuric acid, sulphur is first burned to produce sulphur dioxide. Sulphur dioxide is then oxidized with oxygen to form sulphur trioxide over a vanadium oxide catalyst. Finally, the sulphur trioxide is contacted with water in a highly concentrated sulfuric acid to produce more sulphuric acid at a concentration of 98-99%. The critical part of this process is the oxidation of sulphur dioxide to sulphur trioxide, which is a very exothermic equilibrium reaction that is normally performed in a reactor consisting of four adiabatic catalytic packed beds with intercoolers between beds to reduce the temperature of the gas stream prior to entering the next catalytic bed. The typical variation of the conversion and the temperature through the four catalytic beds is presented in Figure 2-4. The second case study considers the optimization of the first catalytic bed.

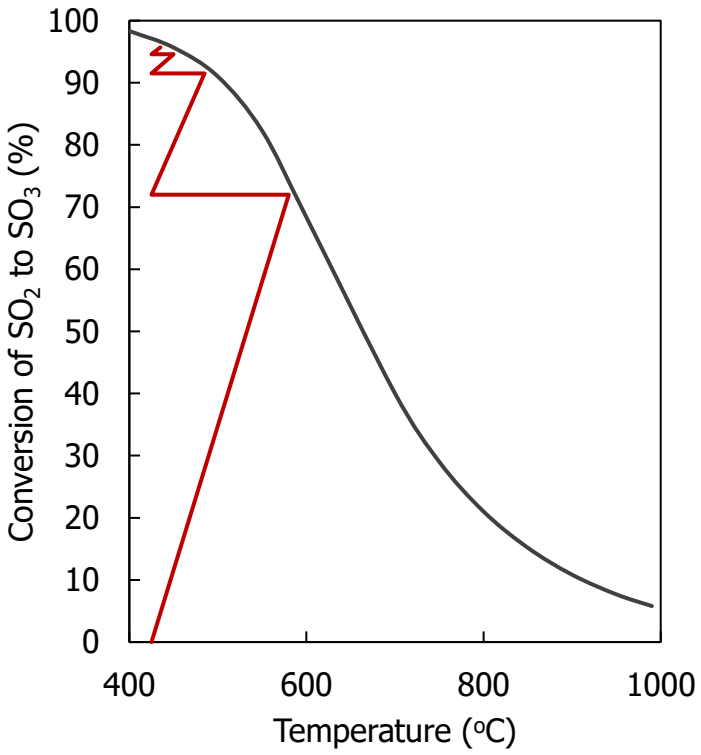

Figure 2-4. Conversion and the temperature profiles through the four catalytic beds with intercoolers between two subsequent beds.

The multi-objective optimization problem consists of two decision variables, namely the inlet gas feed temperature ($T_1 \in$ [500-900] K) and the length of the bed ($L_1 \in$ [0.01-2.00] m), and of three objective functions, namely the SO_3 productivity (P), the SO_2 conversion (X) and the mass of the catalyst (W). The first two objective functions need to be maximized and the third one needs to be minimized. The Pareto domain of this problem was circumscribed using the dual population evolutionary algorithm (DPEA) by varying the two decision variables but fixing many other simulation parameters. The main parameters that were fixed are: bed diameter of 12 m, inlet pressure of 120 kPa, inlet gas molar rate of 800 mol/s with mole fractions of SO_2 of 0.10 and O_2 of 0.11, and a bulk density of catalyst of 256 kg/m^3. The kinetic model of Froment and Bischoff [84] was used for these simulations to predict the reaction rate along the bed. It is further assumed that the process operates under steady-state.

The ranked Pareto domain for the first catalytic bed is presented in Figure 2-5. Instead of having a unique optimal solution, the numerous Pareto-optimal solutions offer

the opportunity to understand the underlying interrelationships that exist between the numerous process variables. Figure 2-5(a) shows the plot of the two decision variables, i.e. the bed length as a function of the inlet gas temperature. Because the reaction is an exothermic equilibrium reaction, if the inlet gas temperature is high, the reaction will tend to reach equilibrium rapidly with a relatively low conversion such that the length of the bed to achieve near-equilibrium will be short. On the other hand, if the inlet gas temperature is low, the length of the bed to achieve near-equilibrium will be much longer and the conversion much higher as shown in Figure 2- 4. The nonlinear relationship observed in Figure 2-5(a) is due the speed of reaction being higher for a higher temperature and the shape of the equilibrium conversion curve.

Figures 2-5(b)- 2-5(d) present the graphs of the three objective functions of the ranked Pareto domain. Figure 2-5(b) clearly shows the typical compromise existing between conversion and productivity. Conversion and productivity have a negative nearly linear correlation. If a high conversion is desired at the exit of the first catalytic bed, then one must accept a low productivity. Conversely, a high productivity can be achieved at the expense of a low conversion. The best-ranked solution has a high conversion with relatively low productivity.

Observing the shape of the curves of Figures 2-5(c) and 2-5(d), one could intuitively select the best ranked solution somewhere around the elbow of the curves. At lower values of the mass of catalyst, the productivity is increased. The higher productivity is associated with a high inlet gas temperature and low catalytic bed length, which obviously lead to a low conversion because near-equilibrium conversion is reached rapidly. This optimization problem is a perfect example to show the different trade-offs that may exist between the different objectives. A method such as Net Flow proves to be very helpful in determining the best ranked solution even though for this simple case study, the decision maker could have probably made his choice based on the graphical representation of the process variables. In this case study, a high conversion and low mass of catalyst were favoured at the expense of a low productivity. However, it is important to recall that the best ranked solution could vary depending on the decision

maker preferences. For this reason, having a full insight on the underlying relationships among the different process variables is very helpful, and this is only possible when examining the Pareto domain.

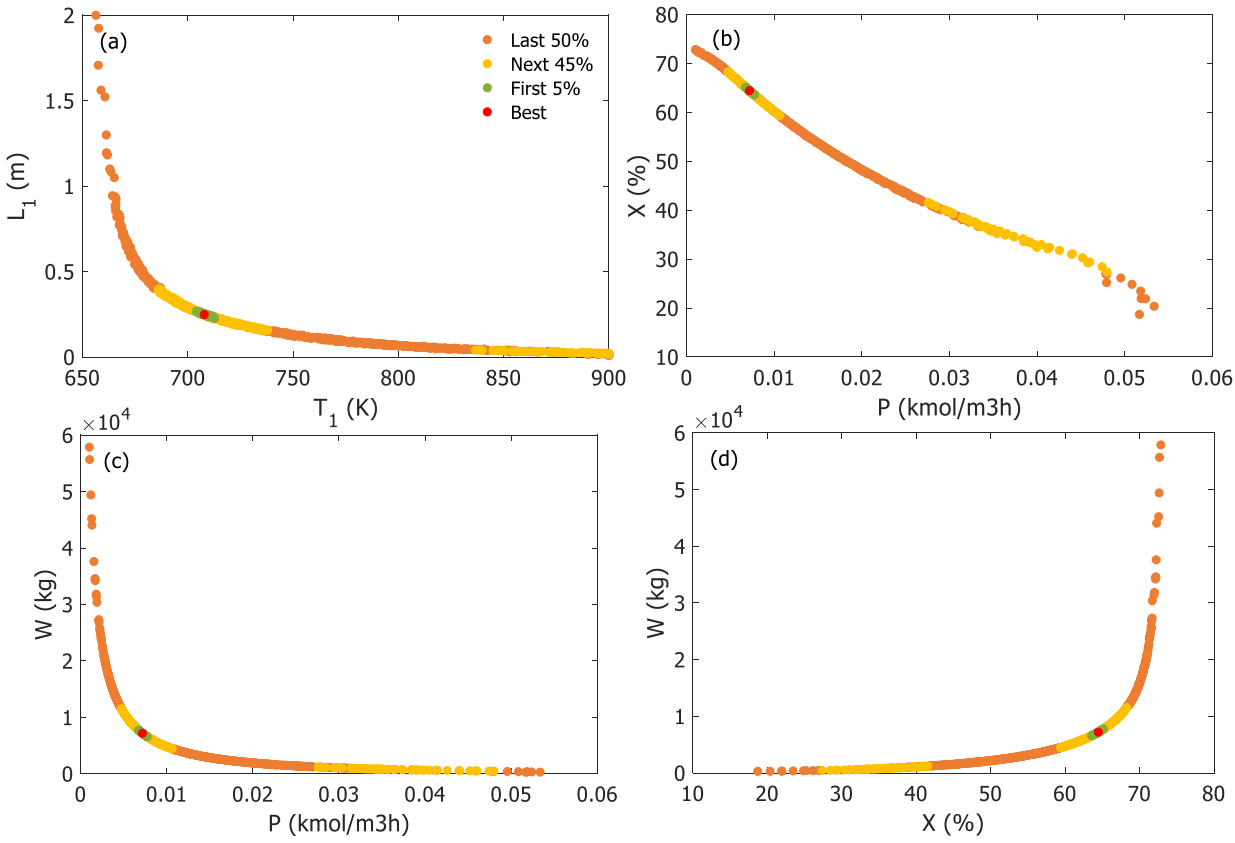

Figure 2-5. Plots of the ranked Pareto domain of the first catalytic bed for the SO_2 oxidation: (a) the two decision variables L_1 and T_1, (b) Conversion as a function of productivity, (c) Mass of catalyst as a function of productivity, and (d) Mass of catalyst as a function of conversion.

4.3 Distillation Column

Distillation is by far the most important separation process in the chemical industry due to its broad applicability. The working principle of this separation unit is to recover certain key component from a mixture based generally on their difference in boiling points or more rigorously on their difference in volatility [85]. This third case study will consist on revisiting this well-known design problem for the separation of a mixture of benzene and toluene.

Toluene (T) is an aromatic hydrocarbon compound and it is often present in paints, lacquers, glues, and solvents [86]. On the other hand, benzene (B) is an organic chemical compound that is primarily used as a feedstock in the manufacturing of ethylbenzene, cumene, cyclohexane, etc. [87] For this case study, the mixture to be separated is an equimolar liquid stream of benzene and toluene. The liquid stream enters the distillation column at its boiling point with a molar flowrate of 100 kmol/h. It is desired to achieve 98 mol% of benzene in the distillate and 98 mol% of toluene in the bottoms product, as shown in Figure 2-6. The design of this distillation column was performed analytically using the Fenske-Underwood-Gilliland (FUG) method and all the main equations are summarized in Appendix A of this chapter.

The relationship between the reflux ratio (*R*) and the number of trays (*N*) is well-known in industry. This information was obtained based on the design equations of a distillation column and trial-and-error experiments. However, other methods could be used to obtain this information in a more explicit way. Multi-objective optimization turns out to be the most appropriate way to achieve this information.

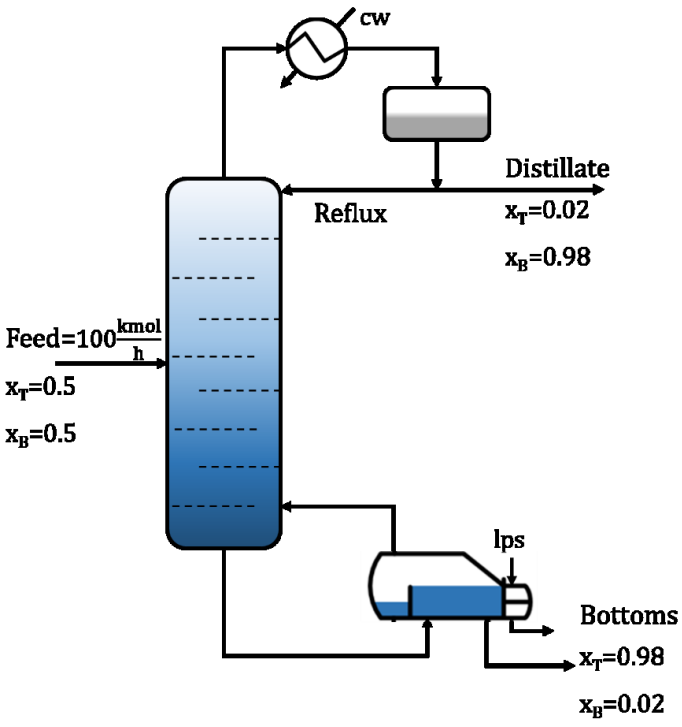

Figure 2-6. Distillation column design for the separation of a toluene-benzene system.

Only one decision variable was considered for this problem, namely the ratio (C) of the actual reflux ratio (R) to the minimum reflux ratio (R_{min}). This decision variable was allowed to vary between 1 and 10. The main reason for carrying an optimization process for a distillation column is to simultaneously minimize two objective functions: the capital ($CAPEX$) and operating ($OPEX$) costs. The capital cost includes the bare module cost (C_{BM}) of the column and its internal parts, as well as the bare module costs of the reboiler and the condenser as shown by Equation (8). As for the operating cost, it considers the cost of steam (13.28$/GJ) and cooling water (0.354$/GJ) used in the reboiler and the condenser, respectively [88].

$$\mathrm{CAPEX} = \mathrm{C_{BM,Col}} + \mathrm{C_{BM,Int}} + \mathrm{C_{BM,Cond}} + \mathrm{C_{BM,Reb}} \quad (8)$$

$$\mathrm{OPEX} = \mathrm{Cost_{cw}} + \mathrm{Cost_{lps}} \quad (9)$$

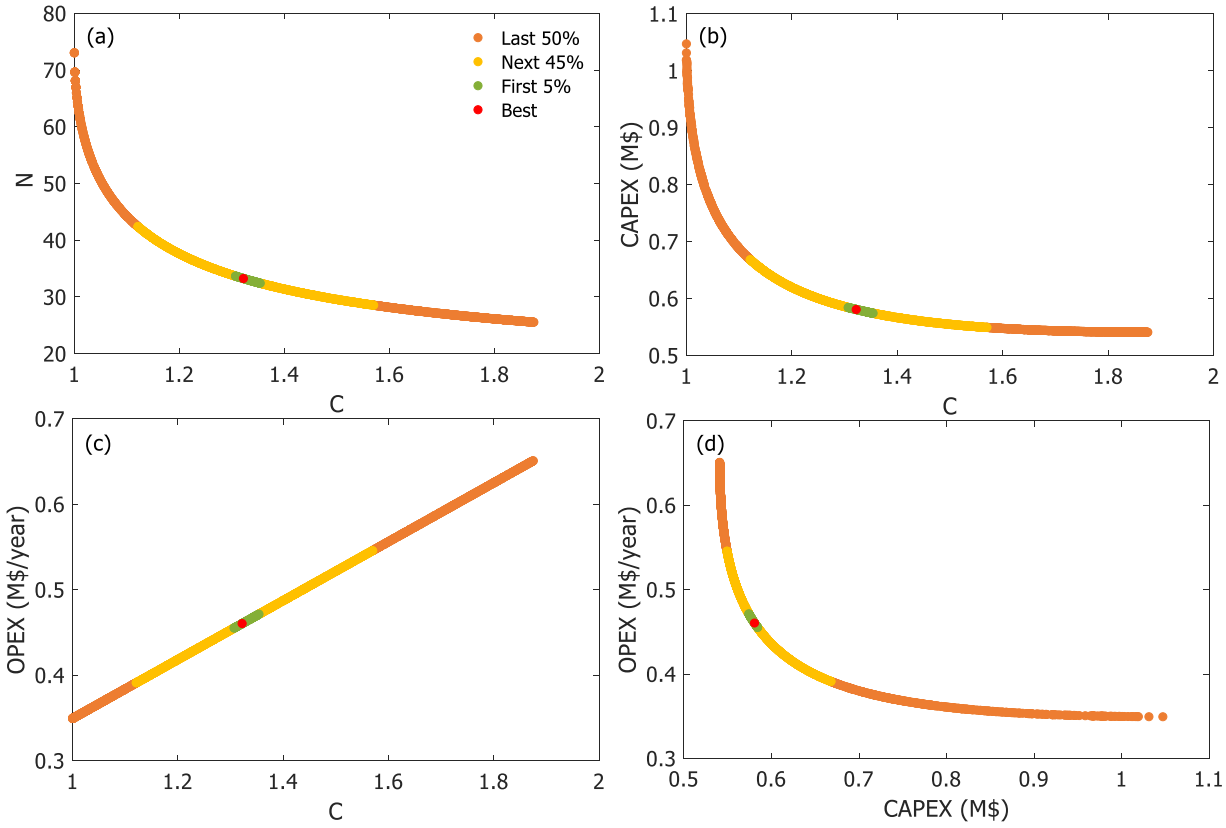

Figure 2-7. Plots of the ranked Pareto domain of the distillation column for the toluene-benzene system: (a) Number of trays against C, (b) Capital cost as a function of C, (c) Operating cost as a function of C, and (d) Operating cost as a function of capital cost.

Figures 2-7(b)-2-7(d) present the ranked Pareto domain. For the ranking of the Pareto domain using Net Flow, a straight line depreciation method over a ten-year period was considered such that the annual capital cost was 10% of the *CAPEX* and such that the minimization problem consists to find the best trade-off between the annual operating and capital costs. Rather than converging to one best solution, which would correspond to the minimum or maximum value of a single objective function lumping the two objectives into one, MOO yields a Pareto domain containing all optimal solutions which allows the DM to choose the best trade-off solution among all the defined and conflicting objectives [89]. From this solution space, it is clear how the selection of factor *C* directly affects the costs of the distillation column. Figures 2-7(b) and 2-7(c) present the relationship between both the *OPEX* and *CAPEX*, and the decision variable. The capital cost has an inverse correlation with the *C* factor so that an increase of this factor incurs in a reduction of the capital cost at the expense of the operating cost which has a positive linear relationship with *C*. The actual number of trays (*N*) was calculated for all different *C* factors and this relation is displayed in Figure 2-7(a). As the number of trays increases, i.e. smaller *C* and reflux ratio, the capital cost increases due to a direct increase in the height of the column, for smaller diameters. However, the non-linear relationship comes from the fact that the capital cost is also dependent on the cost of the reboiler and the condenser that are directly correlated to the energy requirement, and as the *R* value decreases so will the energy usage. Accordingly, the operating cost is reduced since a smaller reflux ratio results in a reduction of the low pressure steam and cooling water usage. On the other hand, as the number of trays decreases which results from larger *C* factor values, the column will be shorter thus reducing the capital cost in general. Although this results in a higher energy requirement as shown by Figure 2-7(c). Again here, by relying on the shape of the curve of Figure 2-7(d), one could intuitively select the best ranked solution somewhere around the elbow of the curve, which corresponds to values of *C* in the range of 1.2 and 1.5. This range for *C* is a well known heuristic in the field of distillation, and the Pareto domain corroborates the range of values of *C*. This shows the utility of obtaining the Pareto domain since many use it without realizing that it contains already known information. It is important to mention that this type of

analysis could be extended for more complex problems, for example for a multi-component distillation where more decision variables will come into play.

4.4 Acrolein Production

Acrylic Acid (AA) plays an important role in the production of polymeric products and a large proportion of its production is converted to a wide range of esters that are applied in surface coatings, textiles, adhesives, paper treatment, polishes, plastics, and many others [90][91][92]. Currently, the preferred production route for AA is by the partial vapour phase catalytic oxidation of propylene in a two-step process, as shown in Figure 2-8. Propylene is first oxidized to acrolein, by supplying a mixture of propylene, air and steam to the first reactor. Acrolein, which is an intermediate product, is subsequently oxidized to AA in the second reactor [93]. For more details on this process model and for all the rate expressions, please refer to Chapter 3 of this thesis and Annex A.

For simplicity, only the first reactor is considered for this case study where a total of 10 chemical reactions, as depicted in Figure 2-9, are involved [94]. An important consideration for this first reactor is the flammability of propylene. For this reason, a constraint on the molar concentration of oxygen [O_2], denoted as Sum_E, was defined and implemented as a soft constraint expressed as an objective function. This objective function represents the minimization of the integration along the length of the reactor of the excess oxygen concentration above the lower flammability limit. The multi-objective optimization problem is then based on three objective functions, namely the maximization of productivity (P) and conversion (X) of propylene to acrolein, and the minimization of the soft oxygen constraint. The set of decision variables consists of five variables and are described in Table 2-2, including their respective lower and upper limits.

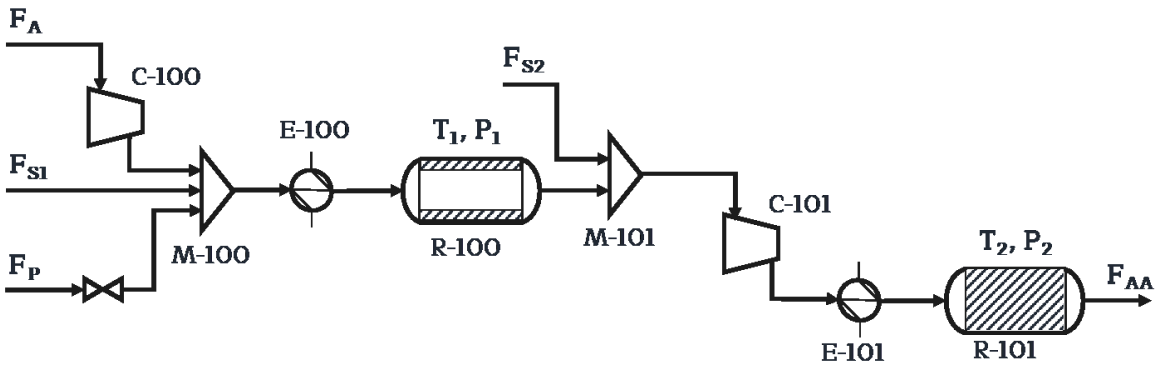

Figure 2-8. Process flow diagram of the two-reactor section of the production of Acrylic Acid.

Table 2-2. Decision variables and allowable ranges for acrolein production.

Decision Variables	x	Min	Max	References
Molar flowrate of propylene [kmol/h]	F_P	91	203	
Molar flowrate of air [kmol/h]	F_A	433	2900	
Molar flowrate of steam [kmol/h]	F_{S1}	91	3047	
Temperature in R-100 [^{o}C]	T_1	330	430	[94][95]
Pressure in R-100 [bar]	P_1	1.05	6	[96]–[99]

It is important to note that at temperatures higher than 360^{o}C, oxygen has a weak influence on the formation of acrolein and the catalytic reduction by propylene will be the rate-determining step. On the other hand, at temperatures lower than 360^{o}C, if the oxygen concentration increases, the formation of acrolein is accelerated. As a result, the expression for the reaction rate from propylene to acrolein will change depending if the temperature is below or above 360°C [94]. To maintain a continuous search space, the rate of reaction was smoothly interpolated between 350 and 365^{o}C using the rate law expressions prevailing before and after the 360°C transition temperature.

Figure 2-9. Reaction scheme for the propylene oxidation reactor [93]***.***

Since a larger number of decision variables are involved in the solution of the multi-objective optimization of the acrolein reactor, the underlying relationships that may exist among the process variables will obviously be more complex. In such cases, it is very useful to have access to the Pareto domain as it will facilitate the task of studying all existing correlations in a more effective way.

Figure 2-10(a) presents the plot of two decision variables, namely the propylene (F_P) and air (F_A) molar flow rates, of the ranked Pareto domain. These two decision variables seem to share a very complex correlation that might stem from the intrinsic constraints within the model of the reactor. This relationship has a direct impact on the productivity and conversion, and Figure 2-10(b) is an example of that relationship. Initially, for values less than 160 kmol/h for F_P, F_P and F_A have a positive correlation as shown by the bands appearing in Figure 2-10(a), which corresponds to the initial portion of the graph. During the transition that occurs around a propylene flow rate of 150 kmol/h, both F_P and F_A have a linear correlation and when the maximum value of propylene flow is reached, there is a drop in the air flow rate resulting in an increase in productivity. This increase is due to the relationships of other process variables with productivity and conversion. For instance, the temperature exerts a strong positive

influence on conversion for temperatures below 430°C (703K), which is the maximum temperature allowed. When the maximum temperature is reached, and a conversion of around 98.7% is achieved, the other four decision variables can push the conversion up to 100%. At the maximum temperature, the rate of conversion of acrolein is only dependent on the propylene partial pressure so that a drop in the air flow rate will increase the partial pressure of propylene resulting in an increase of productivity which is shown on the right side of Figure 2-10(a). Without visualizing the Pareto, it would have been very difficult to infer this information.

It is not shown in Figure 2-10, but the soft constraint on the oxygen concentration is mainly positively dependent on the air flow and more or less dependent on the propylene flow which mean that higher air flow rates will result in the violation of the soft constraint. Furthermore, it is shown that pressure will tend to be around its maximum allowed value since a higher pressure means higher concentrations thus resulting in higher production rates.

In spite of the complex correlations that some decisions variables have with productivity and conversion, these two criteria show a clear, continuous and non-linear negative correlation (see Figure 2-10(d)). In Figure 2-10(d), the small drop observed in the curve around a conversion of 98.7% is due to the strong influence of the temperature when it is lower than its 430°C maximum value. The explanation of the underlying relationships that exist in this model is not complete and the Pareto highlighted the complexity of the underlying relationships that require a deeper analysis and validation of the reactor model. However the purpose of this case study was indeed to show that a complete analysis is only possible due to the availability of the Pareto domain. The Pareto set not only contains all optimal solutions to the optimization process but is also a valuable source of information on the process itself.

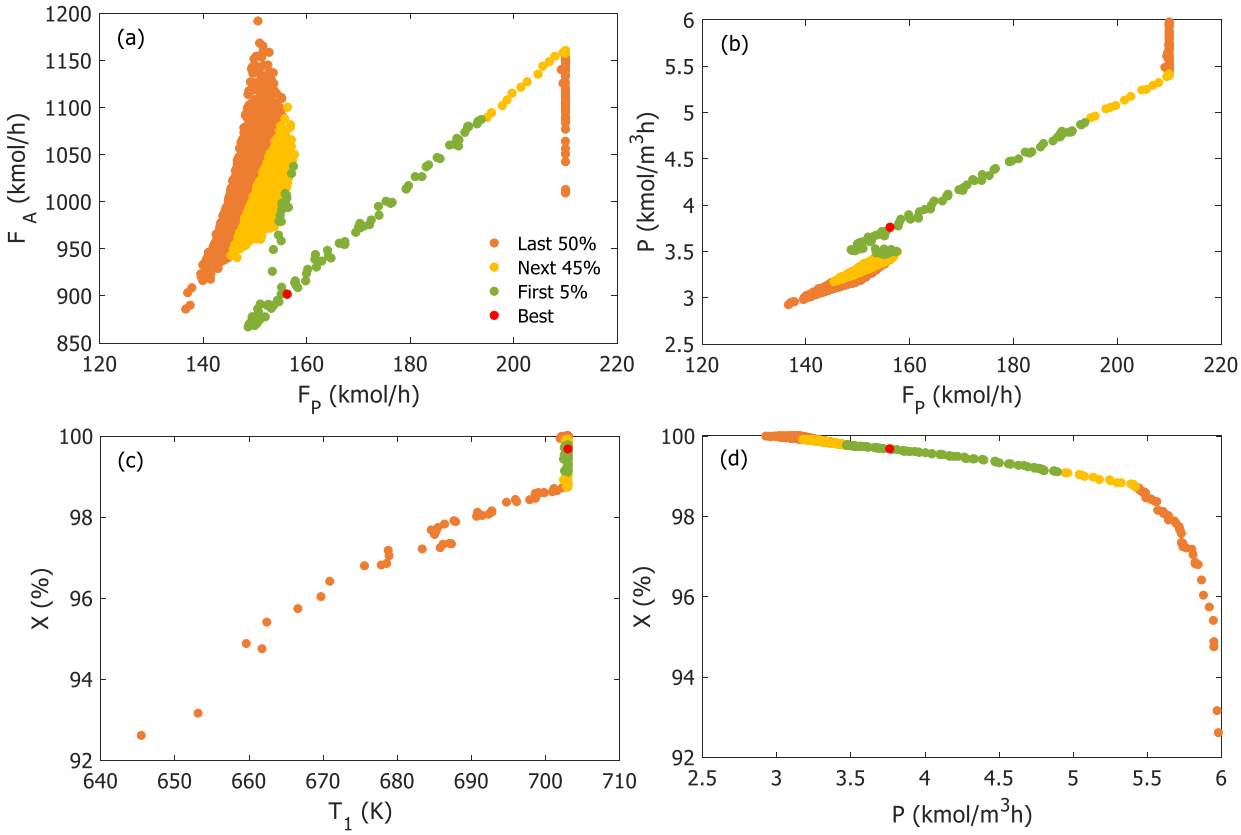

Figure 2-10. Plots of the ranked pareto domain of the acrolein reactor: (a) Air against propylene flow, (b) Productivity as a function of propylene flow, (c) Conversion as a function of T_1, (d) Conversion as a function of productivity.

5. Conclusions

The objective of this paper was to highlight and demonstrate the importance of the Pareto domain in process analysis. In addition of providing the set of optimal solutions to an optimization problem, the solution space is an invaluable source of information that could be used in a myriad of chemical engineering applications. For example, in the case of the PI controller, an in-depth study of the Pareto domain allowed to propose a new controller tuning method where a single variable between the integral time and the controller gain need to be chosen in order to remain optimal. As for the SO_2 to SO_3 reactor, the Pareto domain allowed to have a better insight on the underlying relationships of productivity and conversion as they relate to the length of the bed and the feed temperature. For the distillation column, it was possible to realize how the Pareto domain was in fact embedded into a heuristics to select the optimal reflux ratio

and how it relates to the capital and operating costs. Finally, for the fourth case study, it was shown that the Pareto domain is an invaluable source of information to describe the complexity of the process where a large number of decision variables were considered, to observe the relationships linking all process variables, and to suggest further exploration to gain a better insight in the highlighted complex relationships.

Appendices

Appendix A

The design of the distillation column for the third case study of this investigation was obtained analytically using the Fenske-Underwood-Gilliland (FUG) method. This method consists in finding the minimum number of trays as well as the minimum reflux ratio for a given feed stream, which are then used to determine the other design parameters.

$$N_{min} = \frac{\log\left[\left(\frac{x_{LK}}{x_{HK}}\right)_D\left(\frac{x_{HK}}{x_{LK}}\right)_B\right]}{\log \alpha_{LK/HK}^{ave}} \qquad A\ 2\text{-}1$$

$$\alpha_{LK/HK}^{ave} = (\alpha_{LK/HK}^{D}\alpha_{LK/HK}^{B})^{1/2} \qquad A\ 2\text{-}2$$

$$\alpha_{i/j} = \frac{\frac{y_i}{y_j}}{\frac{x_i}{x_j}} \qquad A\ 2\text{-}3$$

$$\sum_{i=1}^{n} \frac{\alpha_{i/HK} x_{i,F}}{\alpha_{i/HK} - \theta} = 1 - q \qquad A\ 2\text{-}4$$

$$R_{min} = \sum_{i=1}^{n} \frac{\alpha_{i/HK} x_{i,D}}{\alpha_{i/HK} - \theta} - 1 \qquad A\ 2\text{-}5$$

$$R = C \cdot R_{min} \qquad A\ 2\text{-}6$$

$$\frac{N - N_{min}}{N + 1} = 0.75\left[1 - \left(\frac{R - R_{min}}{R + 1}\right)^{0.566}\right] \qquad A\ 2\text{-}7$$

$$E_o = 50.3(\alpha_{LK/HK,F} \cdot \mu)^{-0.226} \qquad \textit{A 2-8}$$

$$N_{real} = \frac{N}{E_o} \qquad \textit{A 2-9}$$

$$N_{actual} = \frac{N}{E_o} + 1 \qquad \textit{A 2-10}$$

$$u_f = C_{sb} F_{ST} F_F F_{HA} \sqrt{\frac{\rho_L - \rho_G}{\rho_G}} \qquad \textit{A 2-11}$$

$$F_{ST} = \left(\frac{\sigma_L}{20}\right)^{0.20} \qquad \textit{A 2-12}$$

$$\frac{2A_d}{A_T} = \begin{cases} 0.1 & \text{for } F_{LV} \le 0.1 \\ 0.1 + \dfrac{F_{LV} - 0.1}{9} & \text{for } 0.1 \le F_{LV} \le 1.0 \\ 0.2 & \text{for } F_{LV} \ge 1.0 \end{cases} \qquad \textit{A 2-13}$$

$$D_T = \sqrt{\frac{4V}{(fu_f)\pi\left(1 - \dfrac{2A_d}{A_T}\right)\rho_G}} \qquad \textit{A 2-14}$$

$$H_P = \begin{cases} 0.4 & \text{for } D \le 0.5 \text{ m} \\ 0.5D^{0.3} & \text{for } D \rangle 0.5\text{m} \end{cases} \qquad \textit{A 2-15}$$

$$H_C = (N_{actual} - 1) \cdot H_P + \sum \Delta H \qquad \textit{A 2-16}$$

The following equations were used to determine the bare module cost of the column, the reboiler and the condenser. In Equation A 2-17, variable *A* represents the capacity size parameter that will change depending on the unit. The height is the capacity parameter used for the column itself and the parameters of the equations are function of the column diameter. The capacity parameter of heat exchangers is the heat transfer area that is determined using Equation A 2-21.

$$C_p = 10\,\text{^}\,(K_1 + K_2 \log(A) + K_3 (\log(A))^2) \qquad \textit{A 2-17}$$

$$C_{BM} = C_p (B_1 + B_2 F_M F_P) \qquad \textit{A 2-18}$$

As the bare module cost is calculated for a CEPCI value of 382 [88], unless otherwise specified, it is necessary to correct for time with the latest cost index: a CEPCI of 616.4 [100]. This was done using the following relation:

$$C_{new} = C_{base}\left(\frac{CE_{new}}{CE_{base}}\right) \qquad A\ 2\text{-}19$$

$$(\cdot \qquad A\ 2\text{-}20$$

$$A = \frac{(\cdot}{U\Delta T_{LM}} \qquad A\ 2\text{-}21$$

As for the column internals, valve trays were selected. The CEPCI value that was initially used in this case is 397. The bare module cost equation for the internals of the column is different:

$$C_{BM} = C_P N F_M F_q \qquad A\ 2\text{-}22$$

Table 2-3. Parameters used for the capital cost (CAPEX) calculation [88]**.**

	K_1	K_2	K_3	F_M	F_P	F_q	B_1	B_2
Condenser	3.4338	0.1445	0.1079	1	1	-	1.8	1.5
Reboiler	3.5638	0.1906	0.1107	1	1	-	1.8	1.5
Column	*	*	*	1	1	-	2.5	1.72
Internals	3.3322	0.4838	0.3434	1	1	1	-	-

*These values will depend on the diameter of the column, see Turton et al. [88]

Appendix B

Table 2-4. Decision variables and objective functions for the four case studies.

PI controller			
	Name	**Bounds or type of optimization**	**Definition**
Decision Variables	Controller gain	[0-3.5]	K_C
	Integration time	[0-3.5]	τ_I
Objective Functions	ITAE	Min	
	OZ	Min	
	ISDU	Min	
SO_2 to SO_3 reactor			
Decision Variables	Inlet gas feed temperature [K]	[500-900]	T_1

	Length of the bed [m]	[0.01-2.00]	L_1
Objective Functions	Productivity [kmol/m^{3}h]	Max	P
	Conversion [%]	Max	X
	Catalyst Weight [kg]	Min	W
Distillation Column			
Decision Variables	Ratio of the actual reflux ratio (R) to the minimum reflux ratio (R_{min})	[1-10]	C
Objective Functions	Capital cost [M$]	Min	$CAPEX$
	Operating cost [M$/year]	Min	$OPEX$
Acrolein Reactor			
Decision Variables	Molar flowrate of propylene [kmol/h]	[91-203]	F_P
	Molar flowrate of air [kmol/h]	[433-2900]	F_A
	Molar flowrate of steam [kmol/h]	[91-3047]	F_{S1}
	Temperature in R-100 [^{o}C]	[330-430]	T_1
	Pressure in R-100 [bar]	[1.05-6]	P_1
Objective Functions	Productivity [kmol/m^{3}h]	Max	P
	Conversion [%]	Max	X
	Soft constraint on oxygen concentration	Min	$SumE$

Appendix C

Below are all the required parameters that were used in order to apply the Net Flow method to rank the Pareto domain obtained for each case study as seen in Section 4 of this chapter.

Table 2-5. Relative weight and thresholds values used in the net flow method to classify the pareto domain obtained for the four case studies .

Case Study	Objective	Relative Weight	Thresholds		
			Indifference	Preference	Veto

PI controller	ITAE	0.4	1.5	4.0	8.0
	OZ	0.4	0.5	1.0	3.0
	ISDU	0.2	0.005	0.01	0.03
SO_3 reactor	P	0.4	0.001	0.003	0.007
	X	0.4	1.2	4.0	8.0
	W	0.2	1500	3000	7000
Distillation column	CAPEX	0.5	2000	5000	8000
	OPEX	0.5	6000	9000	12000
Acrolein reactor	P	0.3	0.08	0.15	0.3
	X	0.3	0.2	0.5	1.0
	Sum_E	0.4	0.02	0.05	0.10

References

[1] S. P. Collette Y., "Introduction: multiobjective optimization and domination," in *Multiobjective Optimization*, 1st ed., Berlin: Springer Berlin Heidelberg, 2004, pp. 15–43.

[2] T. Erfani and S. V. Utyuzhnikov, "Directed search domain: A method for even generation of the Pareto frontier in multiobjective optimization," *Eng. Optim.*, vol. 43, no. 5, pp. 467–484, 2011.

[3] T. W. Rondeau and C. W. Bostian, "Cognitive Techniques: Physical and Link Layers," in *Cognitive Radio Technology*, Second., Blacksbur: Bradley Department of Electrical and Computer Engineering Virginia Tech, 2009, pp. 219–268.

[4] T. Keßler, F. Logist, and M. Mangold, "Use of predictor corrector methods for multi-objective optimization of dynamic systems," in *Computer Aided Chemical Engineering*, 2016, vol. 38, pp. 313–318.

[5] S. Shabeer and M. Y. Wang, "Multi-objective optimization of sequential brakeforming processes," *J. Mater. Process. Technol.*, vol. 102, pp. 266–276, 2000.

[6] B. Li, J. Li, K. E. Tang, X. Yao, and K. Tang, "Many-objective evolutionary algorithms: A survey," *ACM Comput. Surv*, vol. 48, no. 13, pp. 1–34, 2015.

[7] M. T. M. Emmerich and A. H. Deutz, "A tutorial on multiobjective optimization: fundamentals and evolutionary methods," *Nat. Comput.*, vol. 17, no. 3, pp. 585–609, 2018.

[8] M. J. Woodruff, P. M. Reed, and T. W. Simpson, "Many objective visual analytics: rethinking the design of complex engineered systems," *Struct. Multidiscip. Optim.*, vol. 48, no. 1, pp. 201–219, 2013.

[9] X. Blasco, J. M. Herrero, J. Sanchis, and M. Martínez, "A new graphical visualization of n-dimensional Pareto front for decision-making in multiobjective optimization," *Inf. Sci. (Ny).*, vol. 178, no. 20, pp. 3908–3924, 2008.

[10] M. Unal, G. P. Warn, and T. W. Simpson, "Quantifying the Shape of a Pareto Front

in Support of Many-Objective Trade Space Exploration," in *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2016.

[11] K. Deb, "Multi-Objective Optimization Using Evolutionary Algorithms: An Introduction," Kanpur, 2011.

[12] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evol. Comput.*, vol. 10, no. 3, pp. 263–282, 2002.

[13] A. Konak, D. W. Coit, and A. E. Smith, "Multi-Objective Optimization Using Genetic Algorithms: A Tutorial," *Reliab. Eng. Syst. Saf.*, vol. 91, pp. 992–1007, 2006.

[14] H. Wang, Y. Jin, and X. Yao, "Diversity Assessment in Many-Objective Optimization," *IEEE Trans. Cybern.*, vol. 47, no. 6, pp. 1510–1522, Jun. 2017.

[15] A. V. Lotov and K. Miettinen, "Visualizing the Pareto Frontier," in *Multiobjective Optimization*, Berlin: Springer Berlin Heidelberg, 2008, pp. 213–243.

[16] T. R. Newman, J. B. Evans, and A. M. Wyglinski, "Reconfiguration, adaptation, and optimization," in *Cognitive Radio Communications and Networks*, Woodhead Publishing Limited, 2010, pp. 177–198.

[17] I. Y. Kim and O. de Weck, "Adaptive Weighted Sum Method for Multiobjective Optimization," *Struct. Multidiscip. Optim.*, pp. 1–13, 2012.

[18] N. Lalonde, "Multiobjective Optimization Algorithm Benchmarking and Design Under Parameter Uncertainty," Queen's University, 2009.

[19] H. A. Eschenauer, "Multicriteria Optimization Techniques for Highly Accurate Focusing Systems," in *Multicriteria Optimization in Engineering and in the Sciences*, Boston, MA: Springer US, 1988, pp. 309–354.

[20] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidiscip. Optim.*, vol. 26, no. 6, pp. 369–395, 2004.

[21] I. Das and J. E. Dennis, "Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems," *SIAM J. Optim.*, vol. 8, no. 3, pp. 631–657, Aug. 1998.

[22] A. Messac and C. A. Mattson, "Normal Constraint Method with Guarantee of Even Representation of Complete Pareto Frontier," *AIAA J.*, vol. 42, no. 10, pp. 2101–2111, Oct. 2004.

[23] A. Messac, "Physical Programming: Effective Optimization for Computational Design," *AIAA J.*, vol. 34, no. 1, pp. 149–158, Jan. 1996.

[24] A. Messac and C. A. Mattson, "Generating Well-Distributed Sets of Pareto Points for Engineering Design Using Physical Programming," *Optim. Eng.*, vol. 3, no. 4, pp. 431–450, 2002.

[25] Y. Collette and P. Siarry, *Multiobjective Optimization: Principles and Case Studies*. Springer-Verlag Berlin Heidelberg, 2004.

[26] O. Kramer, D. E. Ciaurri, and S. Koziel, "Derivative-Free Optimization," in *Computational Optimization, Methods and Algorithms* , Springer-Verlag Berlin Heidelberg, 2011, pp. 61–83.

[27] G. Chiandussi, M. Codegone, S. Ferrero, and F. E. Varesio, *Comparison of multi-objective optimization methodologies for engineering applications*, vol. 63, no. 5. Elsevier Ltd, 2012.

[28] T. Bäck, *Evolutionary Algorithms in Theory and Practice : Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press, 1996.

[29] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, 2001.

[30] O. M. Shir, S. Chen, D. Amid, D. Boaz, and A. Anaby-Tavor, "Pareto Optimization and Tradeoff analysis applied to Meta-learning of Multiple Simulation Criteria," in *2013 Winter Simulations Conference (WSC)*, 2013, p. 12.

[31] C. M. Fonseca and P. J. Fleming, "An Overview of Evolutionary Algorithms in Multiobjective Optimization," *Evol. Comput.*, vol. 3, no. 1, pp. 1–16, Mar. 1994.

[32] W. Jakob and C. Blume, "Pareto optimization or cascaded weighted sum: A comparison of concepts," *Algorithms*, vol. 7, no. 1, pp. 166–185, 2014.

[33] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. USA: University of Michigan Press, 1975.

[34] D. Mokeddem and A. Khellaf, "Pareto-Optimal Solutions for Multicriteria Optimization of a Chemical Engineering Process Using a Diploid Genetic Algorithm," *Int. Trans. Oper. Res.*, vol. 15, no. 1, pp. 51–65, 2008.

[35] O. Kramer, "Genetic Algorithms," in *Genetic Algorithm Essentials*, Springer, Cham, 2017, pp. 11–19.

[36] V. Kachitvichyanukul, "Comparison of Three Evolutionary Algorithms : GA, PSO, and DE," *Ind. Eng. Manag. Syst.*, vol. 11, no. 3, pp. 215–223, 2012.

[37] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942–1948.

[38] S. Das, A. Abraham, and A. Konar, "Particle Swarm Optimization and Differential Evolution Algorithms: Technical Analysis, Applications and Hybridization Perspectives," *Stud. Comput. Intell.*, vol. 116, no. 2008, pp. 1–38, 2008.

[39] R. M. Storn and K. Price, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *J. Glob. Optim.*, vol. 11, no. 4, pp.

341–359, 1997.

[40] Z. Wang and G. P. Rangaiah, “Application and Analysis of Methods for Selecting an Optimal Solution from the Pareto-Optimal Front obtained by Multiobjective Optimization,” *Ind. Eng. Chem. Res.*, vol. 56, no. 2, pp. 560–574, 2017.

[41] C. Kahraman, “Multi-Criteria Decision Making Methods and Fuzzy Sets,” in *Fuzzy Multi-Criteria Decision Making*, Springer, Boston, MA, 2008, pp. 1–18.

[42] C. L. Hwang and K. Yoon, *Multiple attribute decision making, methods and applications : a state-of-the-art survey*. Springer-Verlag, 1981.

[43] V. Balioti, C. Tzimopoulos, and C. Evangelides, “Multi-Criteria Decision Making Using TOPSIS Method Under Fuzzy Environment. Application in Spillway Selection,” *Proceedings*, vol. 2, no. 11, p. 637, 2018.

[44] L. J. Deng, “Introduction to Grey System Theory,” *J. Grey Syst.*, vol. 1, no. 1, pp. 1–24, 1989.

[45] G.-W. Wei, “Gray relational analysis method for intuitionistic fuzzy multiple attribute decision making,” *Expert Syst. Appl.*, vol. 38, pp. 11671–11677, 2011.

[46] S. H. Sahir, R. Rosmawati, and K. Minan, “Simple Additive Weighting Method to Determining Employee Salary Increase Rate,” *IJSRST*, vol. 3, no. 8, pp. 42–48, 2017.

[47] A. Afshari, M. Mojahed, and R. M. Yusuf, “Simple Additive Weighting approach to Personnel Selection problem,” *Int. J. Innov. Manag. Technol.*, vol. 1, no. 5, pp. 511–515, 2010.

[48] J. Thibault, “Net Flow and Rough Sets: Two Methods for Ranking the Pareto Domain,” in *Multi-Objective Optimization—Techniques and Applications in Chemical Engineering.*, World Scientific: Singapore, 2009, pp. 199–246.

[49] H. Halsall-Whitney, D. Taylor, and J. Thibault, “Multicriteria optimization of gluconic acid production using net flow,” *Bioprocess Biosyst. Eng.*, vol. 25, no. 5, pp. 299–307, Mar. 2003.

[50] V. Srinivasan and A. D. Shocker, “Linear programming techniques for multidimensional analysis of preferences,” *Psychometrika*, vol. 38, no. 3, pp. 337–369, Sep. 1973.

[51] H.-C. Xia, D.-F. Li, J.-Y. Zhou, and J.-M. Wang, “Fuzzy LINMAP method for multiattribute decision making under fuzzy environments,” *J. Comput. Syst. Sci.*, vol. 72, no. 4, pp. 741–759, Jun. 2006.

[52] A. Mardani, E. K. Zavadskas, K. Govindan, A. A. Senin, and A. Jusoh, “VIKOR Technique: A Systematic Review of the State of the Art Literature on Methodologies and Applications,” *Sustainability*, vol. 8, no. 1, pp. 1–38, 2016.

[53] S. Vafaeyan and J. Thibault, “Selection of pareto-optimal solutions for process

optimization using rough set method: A new approach," *Comput. Chem. Eng.*, vol. 33, no. 11, pp. 1814–1825, Nov. 2009.

[54] M. Mendoza Luis Fernando, J. L. Perez Escobedo, C. Azzaro-Pantel, L. Pibouleau, S. Domenech, and A. Aguilar-Lasserre, "Selecting the best portfolio alternative from a hybrid multiobjective GA-MCDM approach for New Product Development in the pharmaceutical industry," in *2011 IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making (MDCM)*, 2011, pp. 159–166.

[55] J. Figueira, V. Mousseau, and B. Roy, "ELECTRE Methods," in *Multiple Criteria Decision Analysis: State of the Art Surveys*, New York: Springer, 2005, pp. 133–153.

[56] B. Filipič and T. Tušar, "Visualization of Pareto Front Approximations in Evolutionary Multiobjective Optimization: A Critical Review and the Prosection Method," *IEEE Trans. Evol. Comput.*, vol. 19, no. 2, pp. 225–245, 2015.

[57] S. Chen *et al.*, "Self-Organizing Maps for Multi-Objective Pareto Frontiers," in *2013 IEEE Pacific Visualization Symposium (PacificVis)*, 2013, pp. 153–160.

[58] B. Filipič and T. Tušar, "A taxonomy of methods for visualizing pareto front approximations," in *Proceedings of the Genetic and Evolutionary Computation Conference on - GECCO '18*, 2018, pp. 649–656.

[59] A. Inselberg, *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*. New York: Springer Verlag, 2009.

[60] J. Li, J.-B. Martens, and J. J. van Wijk, "Judging Correlation from Scatterplots and Parallel Coordinate Plots," *Inf. Vis.*, vol. 9, no. 1, pp. 13–30, Jan. 2010.

[61] W. S. Cleveland, *The elements of graphing data*. Wadsworth Advanced Books and Software, 1985.

[62] N. Elmqvist, P. Dragicevic, and J.-D. Fekete, "Rolling the Dice: Multidimensional Visual Exploration using Scatterplot Matrix Navigation," *IEEE Trans. Vis. Comput. Graph.*, vol. 14, no. 6, pp. 1539–1148, Nov. 2008.

[63] R. H. Koochaksaraei, I. R. Meneghini, V. N. Coelho, and F. G. Guimarães, "A new visualization method in many-objective optimization with chord diagram and angular mapping," *Knowledge-Based Syst.*, vol. 138, pp. 134–154, Dec. 2017.

[64] K. Nöh, S. Niedenführ, M. Beyß, and W. Wiechert, "A Pareto approach to resolve the conflict between information gain and experimental costs: Multiple-criteria design of carbon labeling experiments," *PLoS Comput. Biol.*, vol. 14, no. 10, pp. 1–30, 2018.

[65] E. Zio and R. Bazzo, "Level Diagrams Analysis of Pareto Front for Multiobjective System Redundancy Allocation," *Reliab. Eng. Syst. Saf.*, vol. 96, no. 5, pp. 569–580, 2011.

[66] A. Pryke, S. Mostaghim, and A. Nazemi, "Heatmap Visualization of Population Based Multi Objective Algorithms," in *EMO: International Conference on*

Evolutionary Multi-Criterion Optimization, 2007, pp. 361–375.

[67] E. Li *et al.*, "Improved Heatmap Visualization of Pareto-Optimal Set in Multi-Objective Optimization of Defensive Strategy," in *2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 2018, pp. 345–352.

[68] A. V. Lotov, V. A. Bushenkov, and G. K. Kamenev, *Interactive Decision Maps : Approximation and Visualization of Pareto Frontier*. Springer US, 2004.

[69] P.-W. Chiu and C. L. Bloebaum, "Hyper-Radial Visualization (HRV) method with range-based preferences for multi-objective decision making," *Struct. Multidiscip. Optim.*, vol. 40, no. 1–6, pp. 97–115, 2010.

[70] P. Geiger, "Interactive Graphing of a Pareto Domain," University of Ottawa, 2017.

[71] S. Fettaka, Y. P. Gupta, and J. Thibault, "Multiobjective optimization of an industrial styrene reactor using the dual population evolutionary algorithm (DPEA)," *Int. J. Chem. React. Eng.*, vol. 10, no. 1, 2012.

[72] K. J. Åström and T. Hägglund, *PID controllers: Theory, Design, and Tuning*, 2nd ed. International Society for Measurement and Control, 1995.

[73] Control Station Inc., "Common Industrial Applications of PI Control," 2018. [Online]. Available: https://controlstation.com/pi-control/. [Accessed: 05-Aug-2019].

[74] A. Vandervoort, J. Thibault, and Y. Gupta, "New PI Controller Tuning Methods Using Multi-Objective Optimization," in *Multi-Objective Optimization in Chemical Engineering*, A. Pandu Rangaiah, Gade Bonilla-Petriciolet, Ed. Oxford, UK: John Wiley & Sons Ltd, 2013, pp. 479–501.

[75] J. G. Ziegler and N. B. Nichols, "Optimum Settings for Automatic Controllers," *J. Dyn. Syst. Meas. Control*, vol. 115, no. 2B, p. 220, Jun. 1993.

[76] G. . Cohen and G. . Coon, "Theoretical Consideration of Retarded Control," *Trans ASME*, vol. 75, pp. 827–834, 1952.

[77] C. L. Smith and P. W. Murrill, "Controllers - Set Them Right," *Hydrocarb. Process. Pet. Refin.*, vol. 45, no. 2, pp. 105–124, 1966.

[78] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *J. Process Control*, vol. 13, no. 4, pp. 291–309, 2003.

[79] T. Hägglund and K. J. Åström, "Revisiting The Ziegler-Nichols Tuning Rules For Pi Control," *Asian J. Control*, vol. 4, no. 4, pp. 364–380, 2008.

[80] T. Aziz, M. J. Hossain, T. K. Saha, and N. Mithulananthan, "VAR Planning With Tuning of STATCOM in a DG Integrated Industrial System," *IEEE Trans. Power Deliv.*, vol. 28, no. 2, pp. 875–885, 2013.

[81] C. A. Smith and A. B. Corripio, *Principles and practice of automatic process control*, 3rd ed. New York: John Wiley and Sons, 2006.

[82] S. Tavakoli, I. Griffin, and P. J. Fleming, "Multi-objective optimization approach to the PI tuning problem," in *IEEE Congress on Evolutionary Computation*, 2007, pp. 3165–3171.

[83] S. Mohammad and Z. Alireza, "Comparison of PID Controller Tuning Methods," in *8th National Iranian Chemical Engineering Congress*, 2003.

[84] G. F. Froment and K. B. Bischoff, *Chemical reactor analysis and design*, 2nd ed. New York, 1990.

[85] N. R. Council, "Chemical Industry," in *Separation Technologies for the Industries of the Future*, Washington, D.C.: National Academies Press, 1998.

[86] T. J. Eicher, "Chapter 7-Toxic Encephalopathies I: Cortical and Mixed Encephalopathies," in *Clinical Neurotoxicology; Syndromes, Substances, Environments*, W.B. Saunders, 2009, pp. 69–87.

[87] S. K. Tilley and R. C. Fry, *Chapter 6-Priority Environmental Contaminants: Understanding Their Sources of Exposure, Biological Mechanisms, and Impacts on Health*. Academic Press, 2015.

[88] R. Turton, R. C. Bailie, W. B. Whiting, and J. A. Shaeiwitz, *Analysis, Synthesis, and Design of Chemical Processes*, 3rd Editio. Upper Saddle River, N.J: Prentice Hall, 2009.

[89] D. Savic, "Single-objective vs. Multiobjective Optimisation for Integrated Decision Support, In: Integrated Assessment and Decision," in *Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society*, 2002, vol. 1, pp. 7–12.

[90] H. A. Wittcoff, B. G. Reuben, and J. S. Plotkin, *Industrial Organic Chemicals*, 3rd Editio. Hoboken: WILEY, 2013.

[91] X. B. Xu, J. P. Lin, and P. L. Cen, "Advances in the Research and Development of Acrylic Acid Production from Biomass," *Chinese J. Chem. Eng.*, vol. 14, no. 4, pp. 419–427, Aug. 2006.

[92] M. M. Lin, "Selective oxidation of propane to acrylic acid with molecular oxygen," *Appl. Catal. A Gen.*, vol. 207, pp. 1–16, 2001.

[93] J. J. Farmer, "Process for production of acrylic acid," WO2016130993A1, Feb-2016.

[94] H. Redlingshofer, A. Fischer, C. Weckbecker, K. Huthmacher, and G. Emig, "Kinetic Modeling of the Heterogeneously Catalyzed Oxidation of Propene to Acrolein in a Catalytic Wall Reactor," *Ind Eng Chem Res*, vol. 42, pp. 5482–5488, 2003.

[95] H. Redlingsho, O. Kro, W. Bo, K. Huthmacher, and G. Emig, "Catalytic Wall Reactor

as a Tool for Isothermal Investigations in the Heterogeneously Catalyzed Oxidation of Propene to Acrolein," *Ind Eng Chem Res*, vol. 41, pp. 1445–1453, 2002.

[96] G. Bub, J. Mosler, D. Maschmeyer, A. Sabbagh, R. Fornika, and M. Peuckert, "Process for the production of acrylic acid," 7294741B2, 13-Nov-2007.

[97] O. Hideaki Tsuneki, I. Masanori Nonoguchi, and I. Koji Nishi, "Process for producing acrolein, acrylic acid and derivatives thereof," 9422377B2, 2016.

[98] S. A. Setyowati Wibawanta, "Catalytic Partial Oxidation of Propylene for Acrolein Production," Curtin University, 2011.

[99] T. Shiraishi, S. Kishiwada, S. Shimizu, H. Shigern, I. Hiroshi, and N. Yoshihiko, "Catalytic Process for the Preparation of Acrolein," US3970702A, 1976.

[100] C. Engineering, "Economic Indicators," 2019. [Online]. Available: https://www.chemengonline.com/. [Accessed: 25-Aug-2019].

Chapter 3. Methodology for the Multi-Objective Optimization of Acrylic Acid Production Using Neural Networks as Metamodel

Geraldine Cáceres Sepulveda[1], Silvia Ochoa[2] and Jules Thibault[1*]

[1]Department of Chemical and Biological Engineering, University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
[2]SIDCOP Research Group-Departamento de Ingeniería Química
Universidad de Antioquia, Medellín, Colombia

Abstract

Solving first-principles models could be a computationally intensive problem thus making model-based multi-objective optimization (MOO) a time-consuming task. In this work, a methodology to solve multi-objective optimization for a two-reactor system model producing acrylic acid, using artificial neural networks (ANNs) as metamodels, is proposed in an effort to reduce the computational time required to approximate the Pareto domain. The metamodel performance confirmed a good agreement between the experimental data and the model predicted values of the existent relationships between the eight decision variables and the nine performance criteria of the process. Once the metamodel was built, the Pareto domain was circumscribed based on a genetic algorithm (GA) and ranked with the net flow method (NFM). Using the ANN surrogate model, optimization time decreased by a factor of 15.5.

Nomenclature

AA	Acrylic Acid	-
Ac	Acrolein	-
Ace	Acetaldehyde	-
AceA	Acetic acid	-
ANN	Artificial Neural Network	-
Ci	Molar concentration of i	$kmol/m^3cat$
CWR	Catalytic Wall Reactor	-

DOE	Design of Experiments	-
F_A	Molar flowrate of air	kmol/h
F_P	Molar flowrate of propylene	kmol/h
F_{S1}	Molar flowrate of steam	kmol/h
F_{S2}	Molar flowrate of water vapour	kmol/h
LFL	Lower Flammability Limits	vol%
MOC	Minimum Oxygen Concentration	vol%
MOO	Multi-Objective Optimization	-
$\dot{n}$	Molar flowrate	kmol/h
NFM	Net Flow Method	-
OF	Objective Function	-
n_i	Order of reaction j	-
p_i	Partial pressure of i	bar^{nj}
P_1	Pressure of R-100	bar
P_2	Pressure of R-101	bar
PBR	Packed Bed Reactor	-
Prop	Propylene	-
R	Gas constant	J/mol K
R^2	Coefficient of determination	-
T_1	Temperature of R-100	^oC
T_2	Temperature of R-101	^oC
UD	Uniform Design	-
UFL	Upper Flammability Limits	vol%
V	Volume	m^3
W	Catalyst weight	kg

Greek symbols

η	Efficiency of the compressor	%
ξ	Extent of reaction	kmol/h

1. Introduction

Given the highly competitive market of the chemical industries and the very high capital and operating costs of chemical manufacturing plants, it is paramount to constantly determine the optimal operating conditions, while considering economic, environmental and societal constraints. However, formulating and solving an optimization problem in chemical engineering is not a straightforward task because engineers will invariably have to deal with numerous conflicting/competing objectives. Therefore, there is a need of using optimization strategies that will provide the decision-maker with different alternative solutions that accurately reflect the existing underlying relationships between the process variables, which can be achieved via Multi-Objective Optimization (MOO).

One of the most common approaches used for solving MOO problems is to convert them into a single-objective optimization problem, where an aggregated single objective function is built by taking the weighted sum of the different objectives [1]. However, this approach presents important drawbacks, including: i) the optimal solution is a single point inside the feasible region, and ii) this solution is highly sensitive to the selection of the weights assigned to each objective. In some cases, it may not even be possible to aggregate these competing objectives into a single objective since some of the objectives are of qualitative nature and it may be difficult to numerically quantify them. Solving explicitly a MOO problem can be very expensive in terms of the required computational load given the large number of model evaluations that are required to circumscribe the Pareto domain. Usually, one can develop a phenomenological model of the process to be used with the optimization algorithm, which is usually a very complex and time-consuming endeavour, to generate the Pareto-optimal front, which contains all the non-dominated solutions to the overall problem. Solutions are non-dominated when the improvement of any one of the objectives leads to the deterioration of at least another objective [2][3]. One can also resort to a state-of-the-art process simulator to simulate the entire or part of the process while facing the challenge of interfacing the simulation software with the optimization algorithm. In addition, sometimes the computation time

to perform one simulation run can be relatively long such that the time to circumscribe the Pareto can take a significant amount of time for each optimization scenario. To reduce the simulation time and expedite the determination of the Pareto domain for a given optimization scenario, some authors have successfully used metamodels, also known as surrogate models, to represent the underlying relationships existing between the input or decision process variables and the objectives [4]–[11].

Artificial Neural Networks (ANNs) are often used as metamodels due their high plasticity to encapsulate underlying relationships within the process data. In training ANN models, the connection weights between two layers of neurons are adjusted as to minimize the sum of squares of the errors between the experimental and predicted outputs [12]. The main advantage of using ANNs as surrogate models is their capability of effectively modelling very complex nonlinear behavior, even if it includes a large number of variables. When trained, an ANN can be used to perform a large number of simulations and to determine very rapidly the Pareto domain. In this paper, a methodology for solving Multi-Objective Optimization problems using Artificial Neural Networks as metamodels is proposed and applied to a chemical engineering problem, namely the production of acrylic acid.

2. Description of the Acrylic Acid Production

The comprehensive optimization study of a process normally requires having an accurate and representative model. In this investigation, a first-principles based model was developed for the reactor section of an acrylic acid production plant, which was then used for generating the data in order to train and validate an ANN for solving the MOO. The first-principles based model was simulated in FORTRAN and validated successfully by comparing it with simulation results obtained from ASPEN and Honeywell UniSim. All the main considerations that were accounted for are described in the following sections.

2.1 Reactor Model

Acrylic Acid (AA) plays an important role in the production of polymeric products. The worldwide production of acrylic acid reached ~3.9 million metric tons in 2009 [13].

This important chemical is used for the manufacture of superabsorbent polymers (SAPs) which are involved in a variety of applications [14]. A very large proportion of AA is converted to a wide range of esters that are applied in surface coatings, textiles, adhesives, paper treatment, polishes, plastics, and many others [13][15][16]. In a smaller proportion, it is also used for the production of detergents and flocculants [13].

Nowadays, the commercial production of AA comes from the petrochemical industry [15]. Currently, the preferred production route is by the partial vapour phase catalytic oxidation of propylene in a two-step process, as shown in Figure 3-1. In this process, propylene is first oxidized to acrolein, by supplying a mixture of propylene, air and steam to the first reactor. Acrolein, which is an intermediate product, is subsequently oxidized to AA in the second reactor [14] .

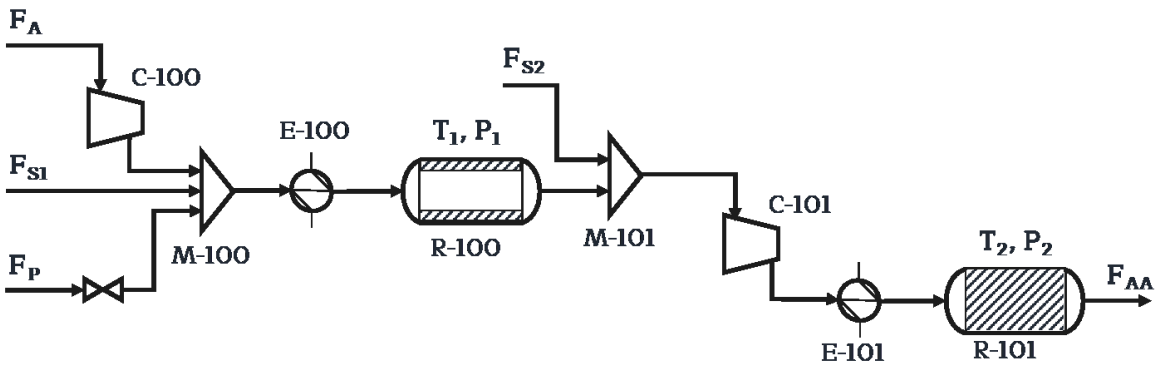

Figure 3-1. Process flow diagram of the two-reactor section of the production of Acrylic Acid.

It is important to note that the desired and undesired reactions are highly exothermic and highly temperature-dependent such that steam, which is fed to the first and second reactors, acts as a thermal sink to moderate the rise in temperature. Furthermore, this process relies on two compressors to bring the reactor feed to the desired operating pressure and, as a result, higher air and water vapour flowrates will significantly increase the compression work. Another important consideration is the flammability of propylene in the first reactor, which will be addressed in the following sections.

2.2 Propylene Oxidation

The propylene vapour phase oxidation occurring in the first reactor is performed in a Catalytic Wall Reactor (CWR), where the catalyst is coated on the inner surface of the reactor and it guarantees isothermal conditions through a temperature control loop for a range of 330-430^{o}C [17]. The oxide catalyst annular section of the CWR is constituted of bismuth and molybdenum containing montmorillonite.

The reaction scheme considered in the model formulation for this reactor is depicted in Figure 3-2, where a total of 10 reactions involving propylene, oxygen, acrolein, acrylic acid, acetaldehyde, acetic acid, formaldehyde, carbon dioxide and carbon monoxide are considered [18]. Given the very low concentrations of formaldehyde observed in the reactor outlet mixture following the series of experiments performed by Redlingshofer et al. [17], it was assumed to be negligible in this study, and was not considered in the model. As it was previously mentioned, steam is present along with propylene and oxygen in the input stream. Steam plays an important role, by increasing the selectivity towards acrolein by supressing the formation of carbon oxides at low temperatures and contributing in the catalyst re-oxidation. In addition, steam has a dilution effect that contributes not only as a thermal sink but also to ensure operating the reactor below the flammability limits [18][19] (more information is provided in Section 2.4 of this chapter).

At temperatures higher than 360^{o}C, oxygen has a weak influence on the formation of acrolein and the catalytic reduction by propylene will be the rate-determining step. On the other hand, at temperatures lower than 360^{o}C, if the oxygen concentration increases, the formation of acrolein is accelerated. As a result, the expression for the reaction rate from propylene to acrolein will change depending if the temperature is below or above 360°C [18].

Figure 3-2. Reaction scheme for the Propylene oxidation reactor [18]***.***

2.3 Acrolein Oxidation

The acrolein vapour phase oxidation kinetics for the second reactor were determined using a catalyst that contains oxides of antimony, nickel and molybdenum [20]. In this case, carbon dioxide was the only major by-product in the reaction scheme. Thus, only the formation of acrylic acid and carbon dioxide were used to determine suitable rate expressions [20].

$$C_3H_4O + \frac{1}{2}O_2 \xrightarrow{r_{10}} C_3H_4O_2$$

R 1

$$C_3H_4O + \frac{7}{2}O_2 \xrightarrow{r_{11}} 3CO_2 + 2H_2O$$

R 2

The experimental study to obtain the kinetics for the acrolein oxidations was carried for a temperature range of 285-315^{o}C [20]. The vapour phase acrolein oxidation reactor was assumed to take place in a Packed Bed Reactor (PBR), and as the catalyst being used is very suitable for acrolein oxidation, the other components aside from steam and air were accounted as inerts.

The complete set of reaction rates considered in the first and second reactors are given in the Appendix section of this chapter.

2.4 Flammability Limits

Operating outside the Lower and Upper Flammability Limits (LFL and UFL) is always recommended to avoid any potentially hazardous situation. Nevertheless, this range will tend to change depending on the temperature, pressure and inert gas concentration inside the reactor. For this reason, it is more convenient to define a Minimum Oxygen Concentration (MOC), below which flame propagation is not possible [21]. Indeed, if the concentration of oxygen is maintained below the MOC, it will be possible to prevent fires or explosions regardless of the concentration of the fuel. The MOC value is expressed in units of volume percent of oxygen in the mixture of fuel and air.

Propylene, which is a flammable gas present in the feed of the first reactor has a LFL and an UFL of 2.4 and 11 vol% in air, respectively, at standard conditions [22]. At 25^{o}C , the MOC in air with nitrogen as the inert gas for propylene as a fuel is 11.5 vol%, while it is 14 vol% when the inert gas is carbon dioxide [21]. The MOC for the maximum possible temperature and pressure of the first reactor was calculated and it was determined that the reactor should be operated below 7.0 vol% of oxygen. As for the second reactor, there is no need to impose a constraint to avoid flammability. If the flammability limit is satisfied in the first reactor, it is expected that it will be satisfied in the second reactor. Furthermore, water is added into the second reactor, which will further dilute the reactor concentration.

3. Methodology for Solving a Multi-Objective Optimization Problem

In this investigation, the production of acrylic acid is optimized using a multi-objective optimization approach. A methodology to solve large multi-objective optimization problems is proposed so that the disadvantage of computational time is addressed when finding the optimal Pareto front. For this purpose, three-layer artificial neural networks are proposed as metamodel to directly predict the objective functions.

The proposed methodology comprises six main steps as illustrated in Figure 3-3. The optimization problem is first established (Step 1) by defining the set of objective functions to be minimized or maximized, and the set of decision variables (with their respective allowable ranges). Building a representative model to perform the optimization study in a timely manner relies on the data set obtained experimentally or through a comprehensive model of the process (Step 2). To obtain the most suitable data, experimental design is usually adopted to gain maximum information with a minimum number of experiments. This is particularly important when the comprehensive model or doing experiments is very time-consuming. In this investigation, Uniform Design (UD) was adopted as a first approach to determine the process data (Step 3) required to build the ANN (Step 4) from the phenomenological model. Once the surrogate model is built, which consists of one ANN for each objective function, it is used to circumscribe the Pareto domain (Step 5) using a genetic algorithm for solving the optimization. In this investigation, a large initial population of 5000 individuals was used. Upon analysis of the prediction performance of the ANNs, possible refinement of the optimization problem is considered such as adjusting the ranges of the decision variables. Finally, all Pareto-optimal solutions are ranked (Step 6) using the Net Flow Method (NFM) where the preferences of a decision maker are embedded in relative weights and three threshold criteria to assist in the ranking of all solutions of the Pareto domain. The optimal solution identified based on the ANN and NFM is then corroborated with an actual experiment or via the comprehensive first-principles based model. More details for some of these steps are addressed in the next subsections, where each step was applied to the optimization of the AA reaction section.

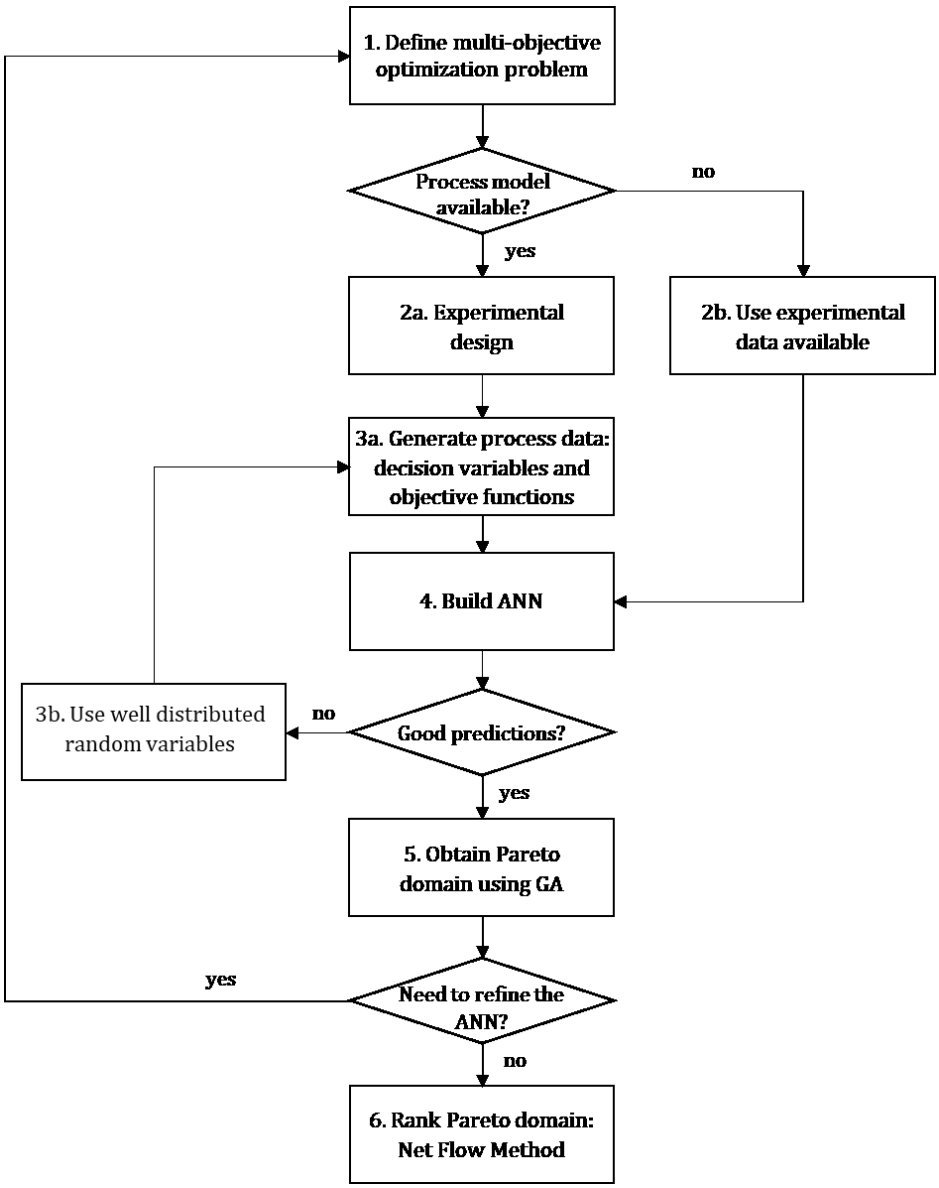

Figure 3-3. Flowchart of the proposed methodology for solving multi-objective optimization using three-layer ANN as metamodel.

3.1 Definition of the Optimization Problem

The multi-objective optimization problem is based on nine objective functions (OF) impacting the economics of the process. These objective functions are presented in Table 3-1 with their respective nomenclature, equation and whether they need to be maximized or minimized. The set of decision or input variables is composed by eight

variables, which are described in Table 3-2, including their respective lower and upper limits.

It is important to notice that the constraint on the molar concentration of oxygen [O_2] in the first reactor, denoted as Sum_E, was implemented as a soft constraint expressed as objective function OF_9. This objective function represents the minimization of the integration of the excess oxygen concentration above the lower flammability limit. The integration for handling the oxygen concentration as a soft constraint allows for an insight on the set of decision variables that violate this constraint and provides more flexibility in the optimization process.

Table 3-1. Objective functions for the MOO of the acrylic acid production process.

Objective Function	Nomenclature	Max or Min?	Equation
Compression Power in C-100 [kW]	OF_1	Min	$\dot{W} = \frac{\dot{n}RT_1\left(\left(\frac{P_2}{P_1}\right)^a - 1\right)}{a \cdot \eta}$
Heat recovery in R-100 [kW]	OF_2	Max	$\dot{H}_{rxn\,j}(T) = \sum_j \Delta H_{rnx\,j}(T) \cdot \xi_j$
Productivity in R-100 [kmol/m^{3}h]	OF_3	Max	$Prod = \frac{F_{Acrolein}}{V}$
Conversion in R-100 [%]	OF_4	Max	$Conv = \frac{(F_{reactant\,in} - F_{reactant\,out}}{F_{reactant\,in}}) \times 100$
Compression Power in C-101 [kW]	OF_5	Min	$\dot{W} = \frac{\dot{n}RT_1\left(\left(\frac{P_2}{P_1}\right)^a - 1\right)}{a \cdot \eta}$
Heat recovery in R-101 [kW]	OF_6	Max	$\dot{H}_{rxn\,j}(T) = \sum_j \Delta H_{rnx\,j}(T) \cdot \xi_j$
Productivity in R-101 [kmol/m^{3}h]	OF_7	Max	$Prod = \frac{F_{AcrylicA}}{V}$
Conversion in R-101 [%]	OF_8	Max	$Conv = \frac{(F_{reactant\,in} - F_{reactant\,out}}{F_{reactant\,in}}) \times 100$

Excess oxygen concentration above LFL	OF_9	Min	$Sum_E=0$ If $[O_2]$ >0.07 then $Sum_E=Sum_E+([O_2]-0.07)*dW$

Table 3-2. Decision variables and their allowable ranges.

Decision Variables	x	Min	Max	References
Molar flowrate of propylene [kmol/h]	F_P	91	203	
Molar flowrate of air [kmol/h]	F_A	433	2900	
Molar flowrate of steam [kmol/h]	F_{S1}	91	3047	
Molar flowrate of water vapour [kmol/h]	F_{S2}	100	4000	
Temperature in R-100 [^{o}C]	T_1	330	430	[17][18]
Temperature in R-101 [^{o}C]	T_2	285	315	[20]
Pressure in R-100 [bar]	P_1	1.05	6	[23]–[26]
Pressure in R-101 [bar]	P_2	3	6	[23][24][26]

The MOO problem can be expressed mathematically by the following equations:

$$\max_{x} OF(x) = (-OF1(x), OF2(x), OF3(x), OF4(x), -OF5(x), OF6(x), OF7(x), OF8(x), -OF9(x))$$

$$\text{Subject to } h_j(x)=0, \quad j=1,\ldots,J$$

$$g_k(x)\leq 0, \quad k=1,\ldots,K$$

$$x_{min} \leq x \leq x_{max}$$

3.2 Design of Experiments

Process data for the production of AA are required to build a representative model that can be rapidly used to generate the Pareto domain of the process. In this study, a set of neural networks, one for each objective function, is used as a metamodel. Historical process data obtained during the actual operation of a process is one source of information. However, for most processes, historical data usually lack generality as they only contain information in a very restricted range of operation. This information is usually not sufficient to build a representative model useful for optimization. To obtain

process data over a wider range of operation, a common approach resorts to experimental design to adequately cover the range of interest for the operation of the process. An experimental design can be used to obtain process data by conducting experiments on the real process or using a comprehensive first-principles based model of the process. An important question arises around the number of experimental points necessary and sufficient to properly develop a representative surrogate model, which in this case means training an ANN for predicting the objective functions based on a set of design points.

Usually, Design of Experiments (DOEs) are used to partition the search domain generated by the independent variables to obtain experimental data that represents the entire domain of interest. The motivation for using DOEs is to determine the underlying relationship that may exist between the decision variables and the objective functions of the process with a restricted amount of data, since a good experimental design should minimize the number of experiments to acquire as much information as possible. Most of the experimental designs assume that the underlying model is known, as it is the case for orthogonal and optimal designs. However, the structure of the model is not always known or may be very complex and highly nonlinear [27]. For such cases, the Uniform Design (UD) proposed by Fang [28] may be used. UD is a design in which the design points are distributed uniformly on the experimental domain to better capture the relationship between the response and the contributing factors [29].

A large number of UD suggested design points have been tabulated and are available online, where each UD takes the form of $U_n(q^s)$ [30]. In this design, a complete table of the normalized design points can be obtained given the dimensionality of factor space *s* (decision variables), the number of levels *q* of the factors, and the desired number of data points *n*. The normalized information obtained from the selected UD table is then used along with the minimum and maximum values of the decision variables to determine the actual values of decision variables from which the objective functions are calculated. The following uniform designs were used to generate the initial set of process data, divided as learning and validation data sets, to build the ANN: $U_{50}(5^8)$ and $U_{20}(5^8)$,

i.e. 50 and 20 design points, which translates to using a ratio of 70/30 for the learning and validation data points.

For more complex cases, where the dimensionality of the input space is large, and a higher number of design points are required in order to develop a good predictive model, one has to resort to another method to define the design points because tables for UD are not available beyond a certain number of design points. In the present investigation, with eight decision variables, it was necessary to use well-distributed random data. This topic will be addressed in more details in the Results section.

3.3 Artificial Neural Networks as Metamodel

Neural networks are very potent metamodels since they have the plasticity to encapsulate the underlying relationships exiting between input and output process variables based on a number of process data. In this work, the simulated data obtained from the phenomenological process model was used to train a three-layer Feedforward Artificial Neural Network (FFANN) for each objective function defined in Section 3.1 of this chapter.

3.3.1 ANN Architecture

Figure 3-4 presents the architecture of the FFANN used in this study, which comprises three layers: the input layer, one hidden layer and an output layer [12]. The input layer is composed of nine neurons: eight neurons corresponding to the decision variables of the MOO problem, and one bias neuron. The only function of the input neurons is to accept the input variables, normalized between 0 and 1, and fan out these scaled input variables to the processing neurons of the hidden layer. The eight decision variables, as defined in Section 3.1, are the molar flowrates of propylene (F_P), air (F_A) and steam (F_{S1}) to the first reactor, the flowrate of water vapour (F_{S2}) to the second reactor, and the operating temperature and pressure of both reactors (T_1, T_2, P_1 and P_2).

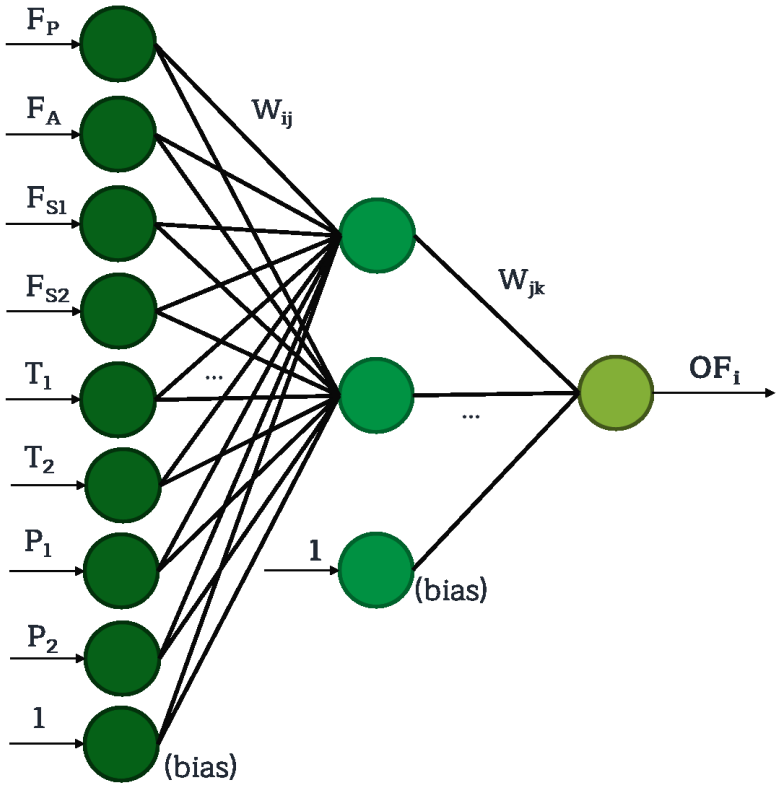

Figure 3-4. Three layer feedforward artificial neural network for the Acrylic Acid MOO problem.

The hidden layer consists of a number of processing neurons and one bias neuron. The number of hidden neurons is chosen as a compromise to obtain an excellent prediction while avoiding overfitting. Each neuron in the hidden layer, except the bias neuron, performs two simple mathematical operations: i) the weighted sum of all outputs of the input layer, including the bias neuron and ii) a nonlinear transformation of the weighted sum using a sigmoid function [11]. Multiple hidden layers could be included, however, it has been shown that a single hidden layer is generally sufficient for classifying most data sets [31]. Finally, the output neuron performs the same two mathematical functions as the neurons of the hidden layer to predict the output variables that are also scaled between 0 and 1. The prediction of the actual output is obtained by de-normalizing the scaled output. For this case study, ANNs with a single output neuron were used to predict independently each objective function to facilitate the learning process of the networks. This implies that the metamodel thereby consists of nine ANNs in parallel to determine each of the nine optimization criteria defined in Section 3.1.

3.3.2 Building the ANN

The experimental data set was divided into learning and validation data to train the networks, and 1000 random test points were used to test each ANN after validation. This "second validation" had zero impact on the ANN, as it was only performed to confirm the good adjustment and precision of the neural networks when exposed to new input data. This additional validation is only possible because a phenomenological model of the process is available. To develop an ANN model of the process based on the series of decision variables and objectives, the connection weights are initially assigned small random values and the predicted output is calculated. The sum of squares of the differences between the actual and predicted output is used to change the connection weights in a way to minimize its value. In this investigation, the quasi-Newton optimization method was used to determine the optimal set of connection weights. The sum of squares based on the learning data set was used to adjust the connection weights until the minimum was achieved, which usually requires a good number of iterations. Throughout the learning process, the sum of squares of the errors based on the validation data set was also calculated at each iteration and the selected set of weights was the one associated with this minimum sum of squares. This procedure is an effective way to avoid overfitting and to choose the adequate number of neurons in the hidden layer [32]. Consequently, the R^2 value was plotted against the number of hidden neurons to determine the proper number of hidden neurons for every ANN. The program used to build the networks was coded in FORTRAN.

3.3.3 Modified Garson Algorithm

Being a black box model, one would usually expect that no information can be retrieved from the surrogate model besides its predictive ability for a set of input data. Nevertheless, knowing the connection weights of the ANNs, it is possible to determine the percentages of the relative importance of each input variable used to generate a specific output in order to perform a sensitivity analysis. For this purpose, the modified Garson method proposed by Goh [33] was used. The algorithm is described by Equation (1): [33]–[35]

$$Q_{ik} = \frac{\dfrac{\sum_{j=1}^{L}\left|W_{ij}W_{jk}\right|}{\sum_{r=1}^{N}\left|W_{rj}\right|}}{\sum_{i=1}^{N}\sum_{j=1}^{L}\left(\dfrac{\left|W_{ij}W_{jk}\right|}{\sum_{r=1}^{N}\left|W_{rj}\right|}\right)} \quad (1)$$

where Q_{ik} represents the relative influence of the input variable *i* on the output variable *k*. W_{ij} corresponds to the connection weight between the input neuron *i* and the hidden neuron *j*, and W_{jk} to the connection weight between the hidden neuron *j* and the output neuron *k*. *i* and *j* stands for the number of the neurons in the input and the hidden layers, respectively. The term $\sum_{r=1}^{N}\left|W_{rj}\right|$ is the sum of the connection weights between the N input neurons and the hidden neuron *j*. L stands for the number of hidden neurons connected to the output neuron *k*. Equation (1) is a modified Garson algorithm since in order to avoid the counteracting influence due to positive and negatives values of the connection weights, the absolute value of each weight was used [34]. It is important to note that the percentages obtained with this equation are normalized such that the sum of all the input's relative importance for one objective function adds up to 100.

3.4 Optimization Algorithm

To solve the multi-objective optimization problem, gradient-free methods are known to be a good alternative [36]. Evolutionary algorithms that are based on Darwin's theory of survival of the fittest have been widely applied to solve these types of problems [37]. The most well-known and widely used method in this category is the Genetic Algorithms (GA), developed by Holland in the 1970s [38]. One major drawback of these algorithms is that they require thousands of evaluations of the process model to reach the Pareto front, which consists of only non-dominated solutions. For a process model that is computationally extensive, which is common, it may require days to circumscribe the Pareto domain. In those cases, the use of a metamodel or surrogate model is very

advantageous. In this investigation, ANNs are used as surrogate models for each of the objective functions.

Once trained and validated, the nine ANNs, one for each objective, are used within the MOO algorithm. To solve the optimization problem, the Dual Population Evolutionary Algorithm (DPEA), coded in FORTRAN, was used in this work [39][40]. As it is the case for other GAs, this algorithm is based on the evolution of a population of individuals, each of which is a solution to the optimization problem [4]. The initial population was comprised of 5000 sets of solutions that were obtained with different sets of decision variables, each decision variable being randomly generated within the permissible ranges. The solutions are then evaluated in pairs to determine the number of times a given solution was dominated. For the next generation, all currently non-dominated solutions were kept along with a fraction of the least dominated solutions. The other solutions were discarded, and the solutions retained from the previous generation were then used to produce new solutions to bring the population to its original number of individuals. This procedure was repeated until 5000 non-dominated solutions were determined. Using a surrogate model, the MOO converges rapidly, and the 5000 non-dominated solutions lead to a well-defined Pareto domain.

3.5 Ranking the Solutions of the Pareto Domain

The Pareto domain was circumscribed without any bias, that is, with no preferences given to any of the objectives, apart from specifying if a given objective needs to be minimized or maximized. It is obvious that some Pareto-optimal solutions are better than other solutions such that a method is required to rank all Pareto-optimal solutions using some preferences expressed by an expert or decision-maker. In this investigation, the Net Flow Method (NFM) was used. This method requires an expert who has good knowledge of the process and can give an appreciation on the nature of each criteria. This information is expressed for each objective function via four parameters, namely: the relative weight (W_k), the indifference threshold (Q_k), the preference threshold (P_k), and the veto threshold (V_k) [41][42]. Using these quantitative parameters, the NFM performs a pairwise comparison of all Pareto-optimal solutions and attributes

a score to each one, which then allows to rank all the solutions. An interesting feature of this ranking method is its robustness, which means that changes in the weights will not incur in major changes of the optimal zones [41].

4. Results

4.1 Construction of the Metamodel

The initial attempt to develop the nine ANNs, as a surrogate model, to represent each of the nine objective functions was performed with a training and validation data sets that were the design points of the Uniform Design. In this first attempt, 50 and 20 design points were used for the training and validation data sets, respectively. Eight decision variables and five levels were used to separate the ranges of the decision variables, which correspond to $U_{50}(5^8)$ and $U_{20}(5^8)$ for the training and validation data sets, respectively. The coefficients of determination (R^2) for each of the nine objective functions (Table 3-1) are plotted as a function of the number of hidden neurons (Figure 3-5).

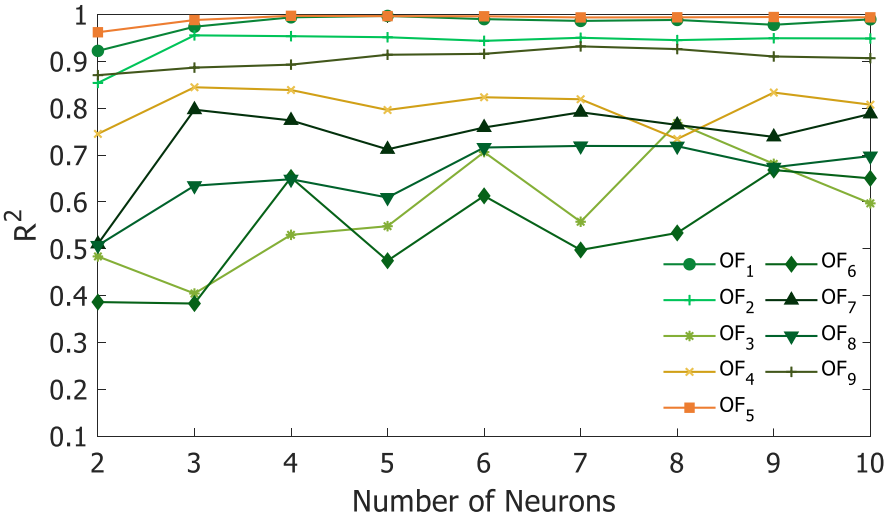

Figure 3-5. R^2 values for all the objective functions using 50 learning and 20 validation UD design points vs. Number of neurons in the hidden layer.

Results of Figure 3-5 show that some objectives, namely the power of the two compressors (OF_1 and OF_5) and the heat recovery of the first reactor (OF_2), are relatively well predicted. However, the ANNs of the other six objectives show poorer predictions

with R^2 values below 0.90. Furthermore, it is not possible to observe a clear trend for those OFs when one would expect the R^2 value to increase as the number of neurons increases. These results suggest that, for these objectives, the number of design data points is insufficient to allow the ANN to capture the underlying relationships that exist between the decision variables and the objectives. The large number of input variables as inputs to the ANNs also points the necessity to present the neural networks with richer information. Since the available tables of Uniform Design are limited to a relatively small number of design points, it was decided to use well distributed random design points, which offer the possibility of using any desired number of design points.

A series of ANNs were developed for an increasing number of hidden neurons and different numbers of design points. The total number of design points were divided in an approximate ratio of 70:30 for the learning and validation data, respectively. Results for objective functions that showed the best and worst predictions, namely the compression power of the first compressor (OF_1) and the heat recovery of the second reactor (OF_6), are presented in Figure 3-6, respectively. The total number of design points (training and validation) in data set varied between 70 and 1430. The predictions for OF_1 are very good for a relatively low number of hidden neurons and a small number of training data points. Indeed, the very high R^2 values indicate that the predictions of the neural network for the compression power of the first compressor are independent of the number of design data points above approximately 140. For the heat recovery of the second reactor (OF_6), the coefficient of determination (R^2) increases with the number of design points whereas it is not a function of the number of hidden neurons above five neurons. This trend was more significant for some objectives due to their dependency on the input or decision variables. For example, a simple dependency prevails for OF_1 as it is mainly correlated to the air flowrate and the operating pressure of the first reactor. In contrast, a much more complex dependency OF_6 as it affected by a larger number of inputs, namely the four input flowrates and the operating temperature and pressure of the first reactor, thereby requiring more data points to capture the underlying relationships between the inputs of the ANN to properly predict this output.

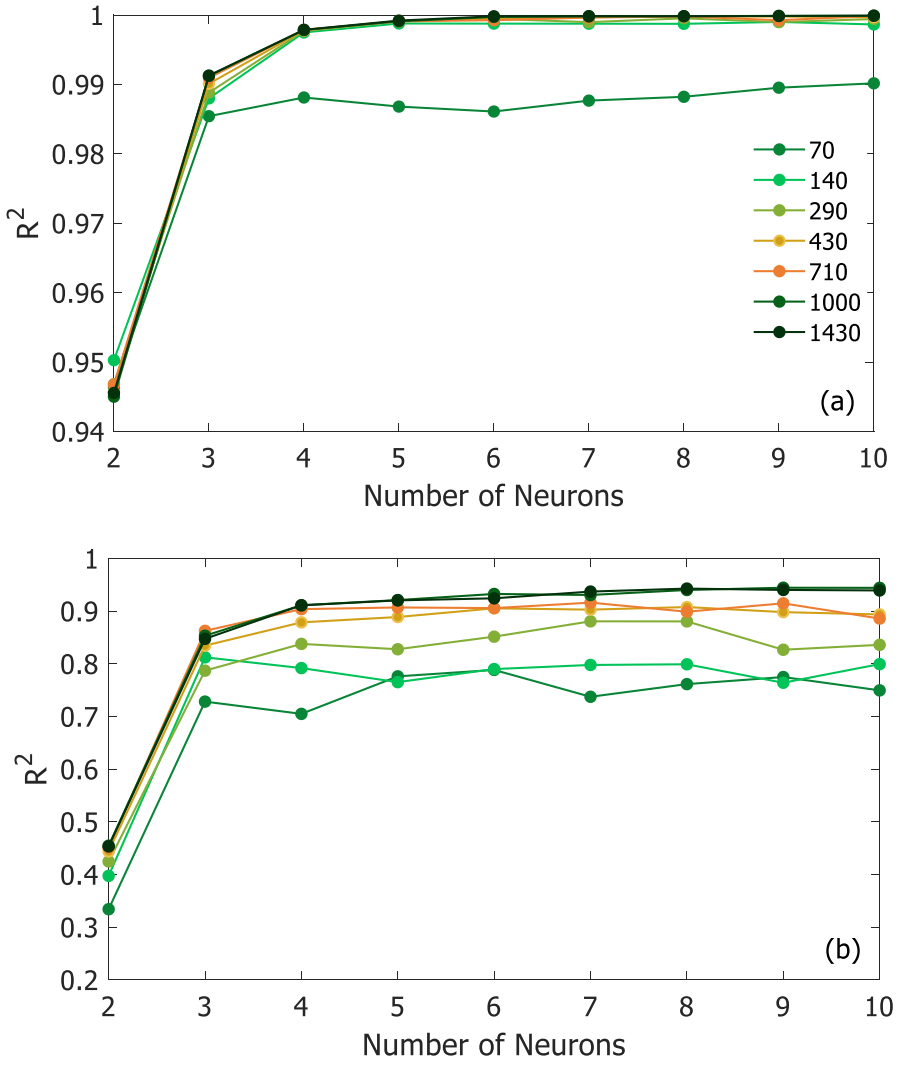

Figure 3-6. R^2 values for different number of random data points used for training and validation vs. Number of neurons in the hidden layer for (a) Compression power in C-100 and (b) Heat recovery in R-101.

Based on the previous discussion, it is interesting to perform a sensitivity analysis to extract the contribution of all neural network inputs to explain each output. Numerous techniques have been proposed to provide this information and partly alleviate the black box character of neural networks. In this investigation, the modified Garson method was used [35]. Results of this sensitivity analysis are presented in Table 3-3 in terms of the percentages of the relative importance of the eight decision variables for each of the objective function in the ANNs. It is important to note that the percentages in Table 3-3 are normalized such that the sum of each row associated to one objective function adds

up to 100; if more variables are correlated to an output, the percentages will be obviously lower. First, these results show the paramount importance of the bias neuron of the input layer, which plays the same role as the intercept of a linear equation to shift the weighted sum to obtain a better fit. Some strong correlations are logically expected and indicate that the neural networks were trained adequately to capture the underlying behaviour of the process. This is the case for instance for the power of the two compressors that are strongly correlated with the desired pressures and the pertinent flow rates. These sensitivity coefficients offer a valuable introspection on the causal effect of each decision variable on the objective functions.

Table 3-3. Relative importance of the input variables on the objective functions in the selected ANN according to the modified Garson method.

	Relative Importance (%)								
Objectives/ Decision Variables	**F_P**	**F_A**	**F_{S1}**	**F_{S2}**	**T_1**	**T_2**	**P_1**	**P_2**	**Bias**
OF_1	1.62	17.14	2.64	1.02	1.76	1.96	41.28	1.79	30.79
OF_2	9.07	13.75	5.23	0.79	27.76	1.93	8.22	1.33	31.90
OF_3	13.42	9.03	3.69	1.24	41.58	0.83	12.25	3.01	14.95
OF_4	5.59	18.22	7.03	4.04	36.30	1.33	19.59	2.01	5.89
OF_5	0.62	11.45	16.24	11.69	6.36	2.22	2.02	26.39	23.02
OF_6	15.24	18.19	9.69	7.66	18.37	1.83	15.42	2.61	10.99
OF_7	17.44	34.58	5.90	6.16	10.72	1.50	8.64	2.80	12.27
OF_8	8.64	38.13	8.07	6.60	13.50	3.42	6.16	7.38	8.10
OF_9	10.91	26.29	15.98	0.86	4.55	0.30	10.09	0.77	30.26

A large number of ANNs to represent the nine objective functions have been obtained. The final selection was made as a compromise of the following criteria: (1) the minimum number of data for training and validating the neural networks, (2) higher than 0.9 values for the coefficient of determination (R^2), and (3) the minimum number of neurons. The selected set of nine ANNs, one for each objective function, can now be used as the surrogate model to generate the Pareto domain and find the optimal operating set of decisions variables.

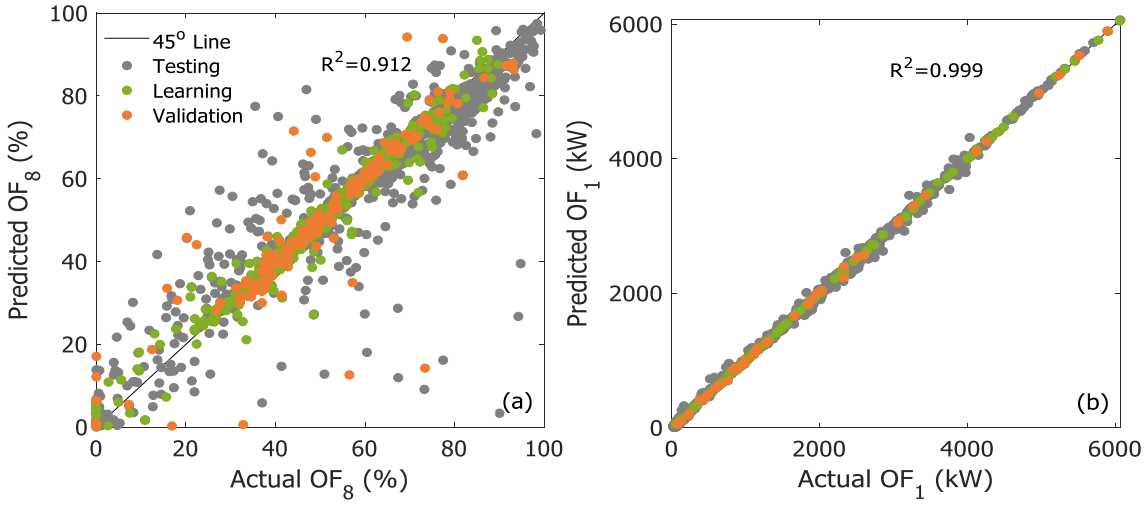

Figure 3-7. Predictions of (a) Conversion in R-101 and (b) Compression power in C-100.

Before proceeding, the quality of the predictions will be examined. Figures 3-7(a) and 3-7(b) present the ANN predictions of the conversion of the second reactor (OF_8) and the compression power of the first compressor (OF_1), respectively, as a function of the values calculated using the phenomenological model. The green points represent the learning data, the orange points correspond to the validation data and the grey points to the testing data. The testing data were generated using the phenomenological model by randomly selecting the decision variables within their allowable ranges, as defined in Table 3-2. This data set is used as a "second validation" to confirm the good adjustment and precision of the ANNs when exposed to new input data. As previously mentioned, it has no impact on the metamodel training. The predictions of Figures 3-7(a) and 3-7(b) correspond to the ANNs with the lowest and highest R^2 values for all data presented: 0.912 and 0.999, respectively. Predictions for the other OFs are very good as well, having R^2 values between the two previous values. The predicted conversion in the second reactor (R-101) has the majority of the points near the 45º-line but with the lowest R^2 value due to few scattered points with poor predictions. When examining the sensitivity parameters of Table 3-3, the two objective functions that are influenced by a larger number of input variables are the conversions of the first and second reactors (OF_4 and OF_8). As mentioned before, the more an objective function is correlated to a larger number of decision variables, the more learning data points are required to obtain better

predictions. As a compromise needs to be done between the R^2 value and the number of learning data, a value of R^2 above 0.9 was considered a good result.

4.2 Multi-objective Optimization

After having obtained a good surrogate model, i.e. consisting of nine AANs, one for each objective function, the Pareto domain was circumscribed with the DPEA and all Pareto-optimal solutions were ranked with the NFM, using both the phenomenological model and the surrogate model for the reactor section to compare the results.

Since there are nine objective functions, the Pareto front is in fact a surface of a nine-dimensional space. To visualize the ranked Pareto domain, it is necessary to resort to two-dimensional projections. In this paper, the ranked Pareto domain of four objective functions are presented for the surrogate and phenomenological models. Figures 3-8(a) and 3-8(b) present the Pareto domain projected on the two-dimensional space of the productivity of acrylic acid (OF_7) and the heat recovery of the second reactor (OF_6), while Figures 3-8(c) and 3-8(d) present the projection on the plane of the conversion of propylene (OF_4) and heat recovery of the first reactor (OF_2). Based on the NFM ranking, the Pareto domain was divided into four different regions: i) the best solution in red, ii) Pareto-optimal solutions ranked in the top 5%, iii) solutions in the next 45% and iv) the remaining 50% of the solutions. The best ranked solution of Figure 3-8(b) correspond to a productivity of 1.179 kmol/m^3h and a heat recovery of 10 755 kW in the second reactor. When the values of the decision variables, associated with Pareto-optimal solution ranked the best, are used within the first-principles based model for comparison purposes, the values for OF_7 and OF_6 are 1.202 kmol/m^3h and 11 104 kW, respectively, yielding errors in the vicinity of 2%-3%. This was also the case for the other objective functions, as shown in Table 3-4. When the phenomenological model is used to circumscribe the Pareto domain and then Pareto-optimal solutions are ranked with NFM, as depicted by Figure 3-8(a), values of 1.2524 kmol/m^3h and 11 291 kW were obtained for OF_7 and OF_6, respectively, for differences of approximately 7% and 5%. The corresponding conversion in R-100 of the best ranked solution is 94.97% and 97.27% for the metamodel and the phenomenological model, respectively, as shown in Figure 3-8(c)

and 3-8(d). In contrast, a conversion of 96.25% was predicted when the decision variables of the best-ranked solution identified with the ANNs were used in the first-principles based model.

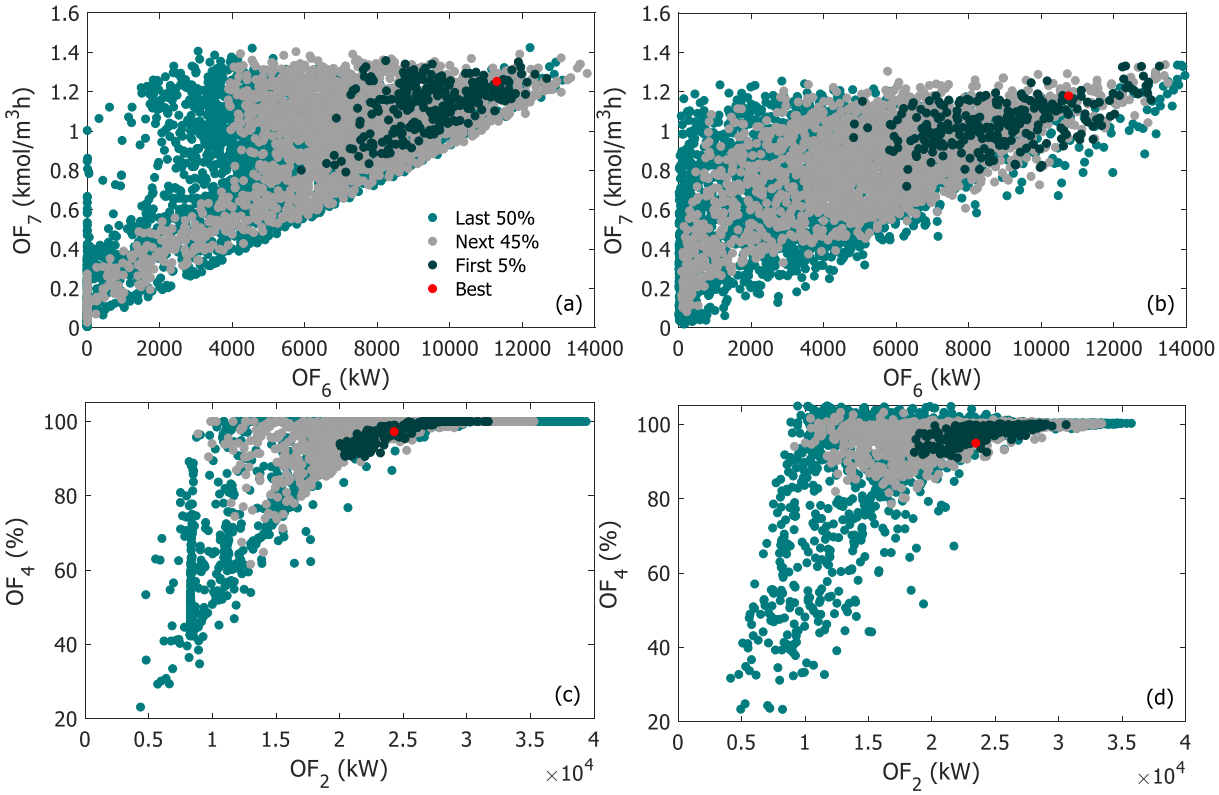

Figure 3-8. Ranked Pareto domain with NFM obtained with: (a) and (c) the phenomenological model and (b) and (d) the ANNs.

The Pareto domain generated with the phenomenological and surrogate models are very similar as illustrated in Figure 3-8. Occasionally, some minor differences between the two Pareto domains will occur. For instance, there is a small region in Figure 3-8(a) that is empty whereas the same region is covered in Figure 3-8(b). Fortunately, it was observed in this investigation that these discrepancies very often appear in regions where Pareto-optimal solutions are ranked relatively low. These discrepancies are usually due to the inability of the neural network to recognize intrinsic constraints embedded in the code of the first-principles based model and that restricts the operation within those limits. As ANN are built from experimental data of the model, it will not be possible for the metamodel to explicitly handle the constraints unless treated as soft constraints as it was the case for the oxygen concentration. In lieu, it filled the empty

region by interpolating the data that is provided to train the ANNs. The best ranked solution as well as the first 5% of the Pareto-optimal solutions were well identified by the metamodel.

Table 3-4. Objective functions of the best ranked solution using NFM from the Pareto domain obtained with a population of 5000; F_P=210.0kmol/h, F_A=1507.9kmol/h, F_{S1}=206.6kmol/h, F_{S2}=100.0kmol/h, T_1=697.65°C, T_2=580.72°C, P_1=1.05bar and P_2=4.01bar.

Objective Function	Metamodel	Phenomenological Model	% Difference
OF_1	90.42	91.10	0.75
OF_2	23460	23700	1.02
OF_3	0.8992	0.9866	9.27
OF_4	94.97	96.25	1.34
OF_5	6581	6794	3.19
OF_6	10755	11104	3.19
OF_7	1.179	1.202	1.93
OF_8	81.67	82.97	1.58
OF_9	0.016	0.000	-

The similarity of the Pareto domains (Figure 3-8) obtained using the phenomenological and surrogate-models is a clear indication that the ANNs were able to adequately predict the existing relationship between the decision variables and the objective functions. To make a more complete comparison between the two Pareto domains, the decision variables and the objective functions of the best ranked solution of the surrogate model were normalized with respect to the best ranked solution obtained with the phenomenological model, where a value of one was assigned to the latter. The normalized variables are presented in Figures 3-9(a) and 3-9(b). These results clearly show that the use of a surrogate model to perform the MOO is a viable solution, as the great majority of the decision variables and objective functions are very closed to the best ranked solution of the phenomenological model. The steam flowrate input is the only variable that has an error above 10%, meaning that using the values obtained from the metamodel will result in 20% more steam usage that what would be required if the first-principles based model was used.

The computation time required to circumscribe the Pareto domain via the surrogate model was 38 s. On the other hand, to obtain the Pareto domain optimizing for the reactor section using the first-principles model took 558 s, which means that the

optimization process was 15.5 times faster using the ANNs. In this particular instance, one simulation with the first-principles model was relatively fast. In other problems, it may take many days of computation time to obtain a sufficient number of Pareto-optimal solutions and this is where the methodology proposed in this work would greatly benefit. In addition, a surrogate model also allows analyzing a large number of optimization scenarios rapidly.

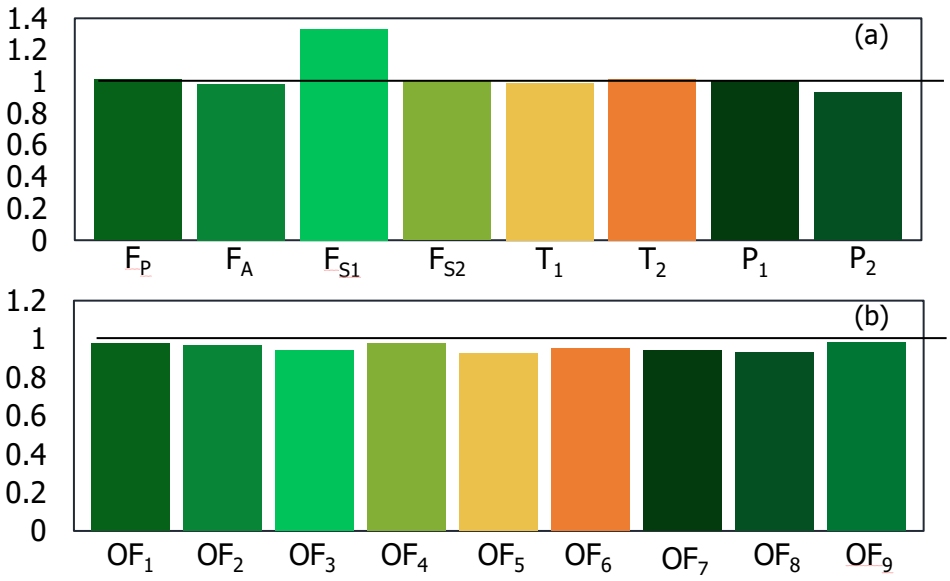

Figure 3-9. Best ranked solution for a) Decision variables and (b) Objective functions using NFM with ANN, normalized with respect to the best solution of the phenomenological model.

5. Conclusions

The aim of this work was to propose an easy-to-follow methodology that would counteract the high computational load of performing multi-objective optimization using first-principles based models. After carrying the optimization process using both, the phenomenological model and the metamodel consisting of nine three-layer ANNs in parallel, one for each objective function, the computational time was reduced by a factor of 15.5 while using the metamodel. This approach can also be very useful in terms of hypothesis testing due to quick convergence of the Pareto domain using ANNs. It should also be noted that the metamodel was able to properly model the existing relationships between the decision variables and the objective functions. This was confirmed by the proper determination of the Pareto domain when comparing the optimization results of

the metamodel against the ones obtained with the mathematical model. Results show that the best ranked solution from the NFM using the ANN metamodel was very close to the optimal solution obtained with the phenomenological model. This work therefore successfully demonstrates the advantage of using ANNs as surrogate model to carry out MOO.

Appendix – Set of rate equations for each reaction

The parameters for the reaction rates involved in the first reactor are presented in Table 3-5 and the rate law equations are presented below.

Table 3-5. Experimental parameters for the rate law of Acrolein formation [18]***.***

	Constants	Value	Units
$\mathbf{r_1}$	$k_{1Red,o}$	0.0628	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$
	$k_{10x,o}$	16000	$\left[\frac{kmol}{kg \cdot s \cdot bar^{0.75}}\right]$
	α_{H2O}	8.2	$\left[\frac{1}{bar}\right]$
$\mathbf{r_2}$	$k_{2,o}$	2.3200	$\left[\frac{kmol}{kg \cdot s \cdot bar^{0.86+0.3}}\right]$
$\mathbf{r_3}$	$k_{3,o}$	0.0150	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$
$\mathbf{r_4}$	$k_{4,o}$	1.4700	$\left[\frac{kmol}{kg \cdot s \cdot bar^{0.73}}\right]$
$\mathbf{r_5}$	$k_{5,o}$	0.0363	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$
	K_{H2O}	1.9	$\left[\frac{1}{bar}\right]$
$\mathbf{r_6}$	$k_{6,o}$	0.00034	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$
	K_{H2O}	1.9	$\left[\frac{1}{bar}\right]$
$\mathbf{r_7}$	$k_{7,o}$	1.3800	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$
	$K_{H2O,AA}$	55.1	$\left[\frac{1}{bar}\right]$
$\mathbf{r_8}$	$k_{8,o}$	0.00038	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$
$\mathbf{r_9}$	$k_{9,o}$	$4.75x10^9$	$\left[\frac{kmol}{kg \cdot s \cdot bar}\right]$

The following equations correspond to the rate law of the acrolein formation (Equations A1 and A2). As previously mentioned, the expression for the reaction rate from propylene to acrolein will change depending if the temperature is below or above

360°C. Below 360°C, if the oxygen concentration increases, the formation of acrolein is accelerated. On the other hand, if the temperature is above 360°C, the catalytic reduction by propylene will be the rate-determining step.

$$\text{If } T_1 \geq 360^{o}\text{C}: \; r_1 = k_{1\text{Red},o} \cdot e^{-\left(\frac{39600}{RT}\right)} \cdot p_{\text{Prop}} \qquad \textit{A 3-1}$$

$$\text{If } T_1 < 360^{o}\text{C}: \; r_1 = k_{1\text{Ox},o} \cdot e^{-\left(\frac{114000}{RT}\right)} \cdot p_{O_2}{}^{0.75} \cdot (2 - e^{-(\alpha_{H2O} \cdot p_{H2O})}) \qquad \textit{A 3-2}$$

The side reactions are described using power-law expressions (A 3-3 to A 3-10). Equations A6-A8 consider water in the adsorption term of the hyperbolic rate expressions since water suppresses the formation of both carbon oxides and leads to higher yields of acrolein and acrylic acid.

$$r_2 = k_{2,o} \cdot e^{-\left(\frac{72500}{RT}\right)} \cdot p_{Ac}{}^{0.86} p_{O_2}{}^{0.30} \qquad \textit{A 3-3}$$

$$r_3 = k_{3,o} \cdot e^{-\left(\frac{52400}{RT}\right)} \cdot p_{O_2} \qquad \textit{A 3-4}$$

$$r_4 = k_{4,o} \cdot e^{-\left(\frac{86700}{RT}\right)} \cdot p_{O_2}{}^{0.73} \qquad \textit{A 3-5}$$

$$r_5 = \frac{k_{5,o} \cdot e^{-\left(\frac{60900}{RT}\right)} \cdot p_{O_2}}{1 + K_{H_2O} \cdot p_{H_2O}} \qquad \textit{A 3-6}$$

$$r_6 = \frac{k_{6,o} \cdot e^{-\left(\frac{38100}{RT}\right)} \cdot p_{\text{Prop}}}{1 + K_{H_2O} \cdot p_{H_2O}} \qquad \textit{A 3-7}$$

$$r_7 = \frac{k_{7,o} \cdot e^{-\left(\frac{82900}{RT}\right)} \cdot p_{O_2}}{1 + K_{H_2O,AA} \cdot p_{H_2O}} \qquad \textit{A 3-8}$$

$$r_8 = k_{8,o} \cdot e^{-\left(\frac{14900}{RT}\right)} \cdot p_{\text{Ace}} \qquad \textit{A 3-9}$$

$$r_9 = k_{9,o} \cdot e^{-\left(\frac{178700}{RT}\right)} \cdot p_{\text{AceA}} \qquad \textit{A 3-10}$$

As for the parameters of the reaction rates involved in the second reactor, they are presented in Table 3-6 and the rate law equations are written below.

Table 3-6. Experimental parameters for the rate law for Acrylic Acid formation [20]***.***

	Constants	Values	Units
$\mathbf{r_{10}}$	$k_{10,o}$	19436	$\left[\frac{1}{s}\right]$
	K_o	$9.78x10^{-6}$	$\left[\frac{1}{s}\right]$
$\mathbf{r_{11}}$	$k_{11,o}$	49070	$\left[\frac{1}{s}\right]$
	K_o	$9.78x10^{-6}$	$\left[\frac{1}{s}\right]$

The following equations correspond to the rate law of the acrylic acid and CO_2 formation:

$$r_{10} = \frac{k_{10,o} \cdot e^{-\left(\frac{55019.6}{RT}\right)} \cdot C_{Ac}}{1 + K_o \cdot e^{-\left(\frac{-31421.84}{RT}\right)} \cdot C_{AA}} \qquad A\ 3\text{-}11$$

$$r_{11} = \frac{k_{11,o} \cdot e^{-\left(\frac{72006.64}{RT}\right)} \cdot C_{Ac}}{1 + K_o \cdot e^{-\left(\frac{-31421.84}{RT}\right)} \cdot C_{AA}} \qquad A\ 3\text{-}12$$

References

[1] R. T. Marler and J. S. Arora, "Survey of multi-objective optimization methods for engineering," *Struct. Multidiscip. Optim.*, vol. 26, no. 6, pp. 369–395, 2004.

[2] T. W. Rondeau and C. W. Bostian, "Cognitive Techniques: Physical and Link Layers," in *Cognitive Radio Technology*, 2nd Ed., Blacksbur: Bradley Department of Electrical and Computer Engineering Virginia Tech, 2009, pp. 219–268.

[3] P. Fleming, "Designing control systems with multiple objectives," in *IEE Master Class. Advances in Control Technology*, 1999, vol. 142, pp. 4–4.

[4] L. Dias, E. Asadi, M. G. da Silva, L. Glicksman, and C. H. Antunes, "Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application," *Energy Build.*, vol. 81, pp. 444–456, 2014.

[5] A. Elmeligy, P. Mehrani, and J. Thibault, “Artificial Neural Networks as Metamodels for the Multiobjective Optimization of Biobutanol Production,” *Appl. Sci.*, vol. 8, no. 6, p. 961, 2018.

[6] F. Farshad, M. Iravaninia, N. Kasiri, T. Mohammadi, and J. Ivakpour, “Separation of toluene/n-heptane mixtures experimental, modeling and optimization,” *Chem. Eng. J.*, vol. 173, no. 1, pp. 11–18, 2011.

[7] C. A. O. Nascimento, R. Giudici, and R. Guardani, “Neural network based approach for optimization of industrial chemical processes,” *Comput. Chem. Eng.*, vol. 24, no. 9–10, pp. 2303–2314, 2000.

[8] L. Magnier and F. Haghighat, “Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network,” *Build. Environ.*, vol. 45, no. 3, pp. 739–746, 2010.

[9] G. A. Tagliarini, J. F. Christ, and E. W. Page, “Optimization Using Neural Networks,” *IEEE Trans. Comput.*, vol. 40, no. 12, pp. 1347–1358, 1991.

[10] R. Altissimi, A. Brambilla, A. Deidda, and D. Semino, “Optimal operation of a separation plant using artificial neural networks,” *Comput. Chem. Eng.*, vol. 22, no. 98, pp. S939–S942, 1998.

[11] H. R. Sant Anna, A. G. Barreto, F. W. Tavares, and M. B. de Souza, “Machine learning model and optimization of a PSA unit for methane-nitrogen separation,” *Comput. Chem. Eng.*, vol. 104, pp. 377–391, 2017.

[12] X. Yao, “Evolving Artificial Neural Networks,” *Proc. IEEE*, vol. 87, no. 9, pp. 1423–1447, 1999.

[13] H. A. Wittcoff, B. G. Reuben, and J. S. Plotkin, *Industrial Organic Chemicals*, 3rd Ed. Hoboken: Wiley, 2013.

[14] “Process for production of acrylic acid,” Feb. 2016.

[15] X. B. Xu, J. P. Lin, and P. L. Cen, “Advances in the Research and Development of

Acrylic Acid Production from Biomass," *Chinese J. Chem. Eng.*, vol. 14, no. 4, pp. 419–427, Aug. 2006.

[16] M. M. Lin, "Selective oxidation of propane to acrylic acid with molecular oxygen," *Appl. Catal. A Gen.*, vol. 207, pp. 1–16, 2001.

[17] H. Redlingsho, O. Kro, W. Bo, K. Huthmacher, and G. Emig, "Catalytic Wall Reactor as a Tool for Isothermal Investigations in the Heterogeneously Catalyzed Oxidation of Propene to Acrolein," *Ind Eng Chem Res*, vol. 41, pp. 1445–1453, 2002.

[18] H. Redlingshofer, A. Fischer, C. Weckbecker, K. Huthmacher, and G. Emig, "Kinetic Modeling of the Heterogeneously Catalyzed Oxidation of Propene to Acrolein in a Catalytic Wall Reactor," *Ind Eng Chem Res*, vol. 42, pp. 5482–5488, 2003.

[19] D. Drysdale, *An introduction to fire dynamics*, 3rd Ed. Wiley, 2011.

[20] V. C. Malshe and S. B. Chandalia, "Vapour Phase Oxidation of Acrolein to Acrylic Acid on Mixed Oxides as Catalyst," *J. appl. Chem. Bioiechnol*, vol. 27, pp. 575–584, 1977.

[21] R. H. Perry and D. W. Green, *Perry's Chemical Engineers' handbook*, 8th Ed. McGraw-Hill, 2007.

[22] M. G. Zabetakis, "Flammability Characteristics of Combustible Gases and Vapors," Washington, 1965.

[23] G. Bub, J. Mosler, D. Maschmeyer, A. Sabbagh, R. Fornika, and M. Peuckert, "Process for the production of acrylic acid," 7294741B2, 13-Nov-2007.

[24] O. Hideaki Tsuneki, I. Masanori Nonoguchi, and I. Koji Nishi, "Process for producing acrolein, acrylic acid and derivatives thereof," 9422377B2, 2016.

[25] S. A. Setyowati Wibawanta, "Catalytic Partial Oxidation of Propylene for Acrolein Production," Curtin University, 2011.

[26] T. Shiraishi, S. Kishiwada, S. Shimizu, H. Shigern, I. Hiroshi, and N. Yoshihiko, "Catalytic Process for the Preparation of Acrolein," US3970702A, 1976.

[27] F. Kai-Tai and D. K. J. Lin, “Uniform Experimental Designs and their Applications in Industry,” *Handb. Stat.*, vol. 22, pp. 131–170, 2003.

[28] F. Kai Tai, “The uniform design: application of number-theoretic methods in experimental design,” *Acta Math. Appl. Sin.*, vol. 3, pp. 363–372, 1980.

[29] R. Li, D. K. J. Lin, and Y. Chen, “Uniform Design: Design, Analysis and Applications,” *Int. J. Mater. Prod. Technol.*, vol. 20, no. 1–3, pp. 101–114, 2005.

[30] K. Yeung, “The Uniform Design,” *Hong Kong Baptist University*, 2004. [Online]. Available: http://www.math.hkbu.edu.hk/UniformDesign/. [Accessed: 12-Mar-2010].

[31] S. Dreiseitl and L. Ohno-Machado, “Methodological Review Logistic regression and artificial neural network classification models: a methodology review,” *J. Biomed. Inform.*, vol. 35, pp. 352–359, 2002.

[32] Y. Liu, J. A. Starzyk, and Z. Zhu, “Optimizing Number of Hidden Neurons in Neural Networks,” in *Proceedings of the 25th IASTED International Multi-Conference: artificial intelligence and applications*, 2007, pp. 121–126.

[33] A. T. C. Goh, “Back-propagation neural networks for modeling complex systems,” *Artif. Intell. Eng.*, vol. 9, no. 3, pp. 143–151, Jan. 1995.

[34] B. Zhou *et al.*, “Relative Importance Analysis of a Refined Multi-parameter Phosphorus Index Employed in a Strongly Agriculturally Influenced Watershed,” *Water Air Soil Pollut*, vol. 226, no. 25, pp. 25–38, 2015.

[35] G. D. Garson, “Interpreting neural-network connection weights,” *AI Expert*, vol. 6, no. 4, pp. 46–51, 1991.

[36] L. T. Biegler and I. E. Grossmann, “Retrospective on optimization,” *Comput. Chem. Eng.*, vol. 28, no. 8, pp. 1169–1192, 2004.

[37] G. Chiandussi, M. Codegone, S. Ferrero, and F. E. Varesio, *Comparison of multi-objective optimization methodologies for engineering applications*, vol. 63, no. 5.

Elsevier, 2012.

[38] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. USA: University of Michigan Press, 1975.

[39] H. Halsall-Whitney and J. Thibault, "Multi-objective optimization for chemical processes and controller design: Approximating and classifying the Pareto domain," *Comput. Chem. Eng.*, vol. 30, no. 6–7, pp. 1155–1168, 2006.

[40] S. Fettaka, Y. P. Gupta, and J. Thibault, "Multiobjective optimization of an industrial styrene reactor using the dual population evolutionary algorithm (DPEA)," *Int. J. Chem. React. Eng.*, vol. 10, no. 1, 2012.

[41] J. Thibault, "Net Flow and Rough Sets: Two Methods for Ranking the Pareto Domain," in *Multi-Objective Optimization—Techniques and Applications in Chemical Engineering.*, World Scientific: Singapore, 2009, pp. 199–246.

[42] A. Vandervoort, J. Thibault, and Y. P. Gupta, "Multi-Objective Optimization of an Ethylene Oxide Reactor Multi-Objective Optimization of an Ethylene Oxide," *Int. J. Chem. React. Eng.*, vol. 9, no. 1, 2011.

Chapter 4. Conclusions and Recommendations

This thesis is mainly focused on two aspects. The first one is emphasizing the importance of carrying multi-objective optimization in order to circumscribe the Pareto domain in an effort to gain a better understanding of the interrelationship of the different variables of a process via the study of several applications from the field of chemical engineering. The second focus of this thesis was to tackle the problem of high computational time when carrying multi-objective optimization using first principles models.

In Chapter 2, the importance of the Pareto domain in process analysis is highlighted and demonstrated through four different chemical engineering applications. More specifically, the design of a PI controller, an SO_2 to SO_3 reactor, a distillation column and an acrolein reactor. For the first application, an in-depth study of the Pareto domain allows to introduce a new controller tuning method where a single variable between the integral time (τ_I) and the controller gain (K_C) needs to be chosen in order to remain optimal as both controller actions (proportional and integral) show a clear linear correlation. For the SO_2 to SO_3 reactor, the Pareto domain allowed to have a better insight on the underlying relationships among all process variables and show that the shape of the Pareto domain could also be used to intuitively select the best ranked solution. As for the third application, it is possible to realize how the Pareto domain is in fact very useful in selecting an optimal reflux ratio in order to reduce the capital and operating costs, and how this information is embedded into a heuristics. Finally, for the acrolein reactor, it was shown that the Pareto domain is an invaluable source of information to describe the complexity of the process where a large number of decision variables were considered, and to observe the relationships linking all process variables.

Then in Chapter 3, an easy-to-follow methodology is developed to counteract the high computational load associated with performing multi-objective optimization using first-principles based models. For this purpose, artificial neural networks are used as metamodels where their main purpose is to be able to approximate each objective

function. After the methodology is applied, the computational time was reduced by a factor of 15.5 while using the metamodel in comparison with the phenomenological model. It is found that this approach can also be very useful in terms of hypothesis testing due to quick convergence of the Pareto domain using ANNs. Results show that the best ranked solution from the Net Flow method using the ANN metamodel was very close to the optimal solution obtained with the phenomenological model which is an indication that the metamodel was well-trained. This work therefore successfully demonstrates the advantage of using ANNs as surrogate models to carry out MOO.

The results from this work highlight the clear benefits from employing multi-objective optimization algorithms in the chemical engineering field. In addition of providing the set of optimal solutions to an optimization problem, the Pareto domain is an invaluable source of information that could be used in a myriad of applications as they allow the decision maker to readily visualize the trade-offs among the different process variables.

In Chapter 3, the uniform design was not applicable as many objective functions were defined for the optimization problem and the restricted uniform points from tables were not sufficient to fully cover the search space. For future work, it would be interesting to determine the maximum number of objective functions that will allow to use the uniform design of experiments. Research was done for steady-state models, were changes in time are not considered. Further research could be done in dynamic multi-objective optimization as the decision variables will also vary according with time.

Annex A. Derivation of the Acrylic Acid Reactor Model Coded in this Work

Detailed explanation of the Acrolein and Acrylic acid reactors

Nomenclature

A	Area	m^2
a	Area exposed to heat transfer per unit volume	m^2/m^3
A_C	Cross sectional area of the reactor	m^2
BBR	Bubbling Bed Reactor	-
C_i	Molar concentration of specie i	mol/L
C_p	Specific molar heat capacity	kJ/mol K
C_{TO}	Initial total concentration	mol/L
D	Diameter	m
D_p	Diameter of the particle in the bed	m
F_j or $\dot{n}$	Molar flowrate of specie j	mol/h
G	Superficial mass velocity	kg/m^2s
g_c	Gravitational constant	-
k	Reaction rate constant	units vary
$\dot{m}$	Mass flow of the reactor fluid	kg/s
N_j	Total number of moles	kmol
P	Pressure	kPa
P_0	Initial pressure	kPa
PBR	Packed Bed Reactor	-
$\dot{Q}$	Rate of heat transfer	kJ/h
R	Gas constant	J/mol K
r_j	Rate of production of j	units vary
T	Temperature	K
t	Time	h
T_0	Initial temperature	K
V	Reactor volume	m^3
W	Catalyst weight	kg
	Shaft work or Power	kJ/h

$\dot{W}_s$	Pressure ratio, P/P_0	-
y	Compressibility factor	-
Z		

Greek symbols

α	Constant dependant on βo, Po and the catalyst properties	-
β_0	Constant dependant on the properties of the PBR	-
$\Delta H°_f$	Enthalpy of formation	kJ/mol
$\Delta H_{rxn\ j}$	Heat of reaction for reaction j	kJ/mol
η	Efficiency of the compressor	%
μ	Viscosity	Pa-s
ξ	Extent of reaction	kmol/h
ρ_0	Density of the reactor fluid	kg/m^3
ρ_b	Bulk density	kg/m^3
ρ_C	Catalyst density	kg/m^3
ϕ	Void fraction	-

A.1 Propylene Oxidation

Modeling of the propylene oxidation reactor can either be made so that a specific amount of catalyst is used or until a specific amount of a certain chemical is reached. In this sense, the design equation for the reactor in terms of the molar flows for a packed-bed reactor (PBR) is:[1]

$$F_{j,o} - F_j + \int_W r_j dW = \frac{dN_j}{dt} \qquad 1$$

Where the o stands for the inlet flow.

The reactor operates at steady state so that the accumulation term will leave. After applying the mole balance equation for a differential catalyst weight, ΔW, dividing by ΔW and taking the limit to zero, the differential expression that will relate the flow of a chemical *j* to the amount of catalyst present in the reactor is as follows:[1]

$$\frac{dF_j}{dW} = r_j \qquad 2$$

The stoichiometric coefficients of each reaction are used to determine the rate of formation of each chemical with respect to the rate law that was given for each reaction, a detailed procedure is given in Annex B. Once all the rates are found it is possible to replace them into the design Equation 2 for each component. In order to solve the differential equations, Euler's method was used with a small step size to ensure accuracy and precision in the results.

An expression to calculate the partial pressure of every component is required to calculate the rate laws. The reactor will be isothermal and isobaric and thus the following equation will be used: [1]

$$P_j = C_{To} \frac{F_j}{F_T} RT \qquad 3$$

Since the rates of reaction $r_{4AceticA}$ and r_{7CO2} are not dependent of the partial pressure of acetaldehyde and acrylic acid respectively, the rates would be calculated even if there were no flows for these chemicals. For this reason, a condition statement was implemented to set the rate equations to zero given the case where the flows of acetaldehyde or acrylic acid were zero. In this order of ideas, another condition needed to be implemented for the flow of acrylic acid. As no acrylic acid is present at the entrance of the reactor, its net rate of formation cannot be negative, therefore if it becomes negative it will be set to zero.

Another important thing to account for, is that one need to be aware of the drawbacks of using a numerical method to solve the reactor model. It is not physically possible to have negative flows. For this reason, when the flow of the reactants becomes negative, it will be set to zero.

A.2 Acrolein Oxidation

In this case, the rate law expressions are dependant of the components concentration. The reactor operates isothermally but will have a pressure drop so that the following equation for concentration is used: [1]

$$C_j = C_{To}\frac{F_j}{F_T}\frac{P}{P_o} \qquad 4$$

As mentioned in the section above, the procedure for the design of this reactor will be very similar to the propylene oxidation reactor. However, compared to the other reactor this one does have a pressure drop and it will be calculated using the Ergun equation:[1]

$$\frac{dy}{dW} = -\frac{\alpha}{2y}\frac{F_T}{F_{To}} \qquad 5$$

Where

$$y = \frac{P}{P_o}$$

$$\alpha = \frac{2\beta_o}{A_c \rho_c (1-\phi) P_o}$$

$$\beta_o = \frac{G(1-\phi)}{\rho_o g_c D_p \phi^3}\left[\frac{150(1-\phi)\mu}{D_p} + 1.75G\right]$$

A.3 Criteria for Both Reactors

The optimization problem has 9 different criteria: 4 for each reactor, namely, maximization of the productivity and the conversion of propylene and acrolein of both reactors. The minimization of the power of the compressor used to compress the air before entering the first reactor and of the compressor required to increase the pressure of the mixture going to the second reactor. Maximization of the heat recovery of both reactors since it could be used for steam generation and the last one will be the

minimization of the restriction on the oxygen concentration, i.e. integration of the excess oxygen concentration above explosive limit in the first reactor.

The productivity of both reactors was calculated in relation to the desired product of each reactor. To do so, the volume of the reactor itself was required.

$$\text{Productivity} = \frac{F_{main\,\text{Product}}}{V} \quad 6$$

$$V = \frac{W}{(1-\phi)\rho_c} \quad 7$$

Additionally, the conversion of each reactor in terms of the main reactant was calculated as follows:

$$\text{Conversion} = \frac{(F_{\text{reactant in feed}} - F_{\text{reactant in outlet}})}{F_{\text{reactant in feed}}} \times 100 \quad 8$$

As for the compression power, $\dot{W}$, the following equation for reversible adiabatic power was used:[2]

$$\dot{W} = \frac{\left(\left(\frac{P_2}{P_1}\right)^a - 1\right)}{a \cdot \eta} \quad 9$$

Where a is a constant value equal to $\frac{k-1}{k}$ and where k stands for C_p/C_v. In this case, the compressibility factor Z is equal to 1 since an ideal gas is assumed. This assumption is possible since it is known that a gas behaves more like an ideal gas at higher temperature and lower pressure.

From heuristics, the efficiency for reciprocating compressors will be 65% at compression ratios of 1.5, 75% at 2 and 80-85% at 3-6 [2]. Using this information, Figure A-1 was plotted to be able to approximate the efficiency of the compressor depending on the compression ratio. The dashed line represents the data given by heuristics and as one can see, every section from Table A-1 was approximated by a line. All the equations are displayed in the figure.

It is important to note that the compression ratio will be determined using the following equation:

$$\text{Compression ratio} = \left(\frac{P_N}{P_1} \right)^{\frac{1}{N}} \qquad 10$$

Where N is the number of stages and it will be determined using the outlet and the inlet pressure of the compressor.

Table A-1. Number of stages based on the outlet and inlet pressure of the compressor.

P_{out}/P_{in}	Number of stages (N)
<4	1
4 to 16	2
16 to 64	3
64 to 256	4

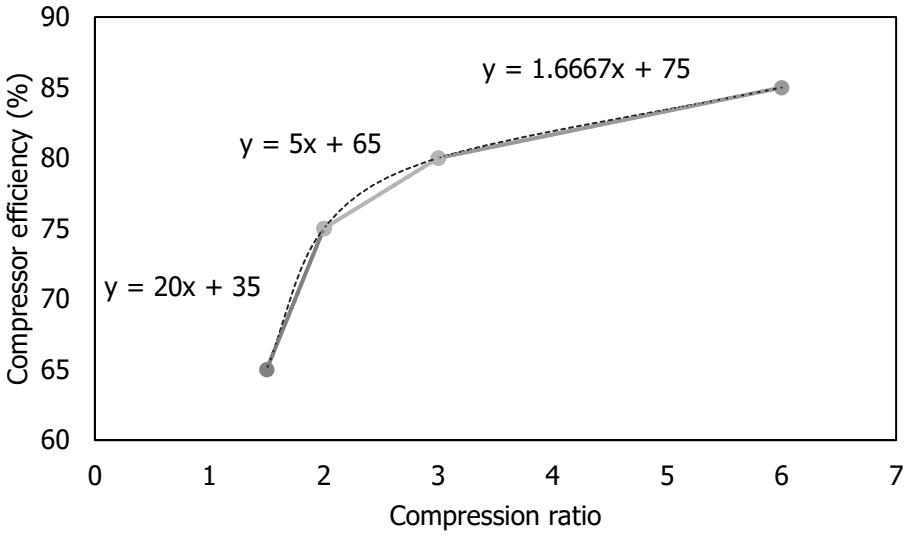

Figure A-1. Determination of the compression efficiency.

From the compression power equation, it is evident that the inlet temperature of each compressor is required. For the first compressor, the air inlet temperature will always be assumed to be 25°C, thus, the k value is assumed constant at 1.398. However, the inlet temperature for the second compressor will always vary depending on the operating temperature of the first reactor and the steam temperature. The total C_p value

of the outlet flow of the first reactor was required. For this purpose, the C_p for each component was calculated at the operating temperature of the reactor and it was then multiplied by the component molar fraction, and the sum of these values yielded the total C_p value. Steam was assumed to enter the process at 100°C, and its C_p value is 34.243 kJ/mol*K at this condition. To obtain the k values for the outlet flow of the first reactor and the steam, the following relation for ideal gas was used:

$$k = \frac{C_p}{C_v} = \frac{C_p}{C_p - 1} \qquad 11$$

To get the k value of the mixture of the steam and the outlet flow of the reactor, each component's k value was multiplied by its molar fraction and then summed. As it was mentioned before, both flows are at different temperatures, so that a mixing temperature was calculated in order to obtain the compression power.

$$T_{mix} = \frac{F_{steam} * C_{p,steam} * 373.15K + F_{outR_1} C_{p,outR_1} T_{outR_1}}{F_{steam} * C_{p,steam} + F_{outR_1} C_{p,outR_1}} \qquad 12$$

It is well known that as the pressure of a gas increases, its temperature will also increase. Nevertheless, a heat exchanger before each reactor was used to take care of this change in temperature and bring the gas to the required operating temperature. The pressure drop of the heat exchangers was assumed negligible for simulation purposes.

Finally, the heat recovery in both reactors was calculated. The importance of this criteria comes from the economic profitability of steam generation. Although, a good control of the process is required to avoid any hazardous situation. The heat recovery is directly related to the total heat of reaction of the reactor. It was assumed that the heat of reaction of both reactors could be completely recovered and that both reactors could operate isothermally along the length of the reactors. Since every reaction will occur at a different rate, the extent of reaction also needed to be accounted for to determine the total heat of reaction released by every reaction.

A material balance was applied to each specie, where the extent of reaction ξ was defined for every reaction *j*. ξ is the moles per unit time of main reactant to react in

reaction *j*. In the same way, the stoichiometric coefficient *ν* (positive for the products and negative for the reactants) was defined for every component *i* involved in reaction *j*.[3] In this case *i*=1-9 and it corresponds to propylene, O_2, H_2O, acrolein, acrylic acid, CO_2, CO, acetaldehyde and acetic acid, respectively.

$$\dot{i} \quad \cdot \quad \xi_{\nu_j} \qquad 13$$
$$j$$

For the first reactor, 9 species were directly involved in Reactions 1-9 which was translated into the following 9x9 matrix.

$$\begin{pmatrix} -1 & 0 & -1 & 0 & -1 & -1 & 0 & 0 & 0 \\ -1 & -0.5 & -2 & -0.5 & -3 & -4.5 & -3 & -2.5 & -2 \\ 1 & 0 & 1 & 0 & 3 & 3 & 2 & 2 & 2 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 3 & 3 & 2 & 2 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \\ \xi_8 \\ \xi_9 \end{pmatrix} = \begin{pmatrix} \dot{i} & \cdot \\ \dot{i} & \cdot \\ \dot{i} & \cdot \\ \dot{i} & \cdot \\ \dot{i} & \cdot \\ \dot{i} & \cdot \\ \dot{i} & \cdot \\ \dot{i}_{,out} - \dot{i} & \\ \dot{i} & \cdot \end{pmatrix}$$

To find the extent of reaction, $\xi_{1\text{-}9}$, and then find the total rate of heat of reaction, it was necessary to solve the matrix. Nevertheless, this is a singular matrix and methods requiring the inverse of the matrix could not be used to solve it. To ensure that the answer was appropriate, a different approach was used to solve for the 9 extent of reaction.

For this purpose, the VF04AD subroutine was implemented. This subroutine calculates a local minimum for an objective function subject to nonlinear equality and inequality constraints:[4]

$$\min f(x_1, x_2, ..., x_n)$$
$$s.to$$
$$g_i(x_1, x_2, ..., x_n) = 0$$
$$h_j(x_1, x_2, ..., x_m) \geq 0$$
$$i = 1, ..., n$$
$$j = 1, ..., m$$

The subroutine will also be able to handle linear constraints. They will pass as "nonlinear" constraints into the subroutine.

In this case, the constraints will be defined based on the material balance of each specie given by Equation 12, except that the $\dot{n}_{i,out}$ term will be sent to the other side of the equation, so that they will be all linear equality constraints. The objective functions will correspond to the sum of all the constraints equations which should be equal to zero. The optimization will then yield the values of the ξ parameters which will satisfy the problem.

This procedure will not be necessary for the second reactor since there are only 2 reactions inside the reactor and one will be able to directly solve for ξ_1 and ξ_2. After applying Equation 13 for all the species in the reactor, the following equations were used to calculate the extent of reaction for both reactions of this reactor.

$$\xi_1 = i \qquad 14$$

$$\xi_2 = \frac{j}{3} \qquad 15$$

After the extent of reaction was calculated for all the reactions, it was possible to proceed with the calculation of the heat of reaction for every reaction *j*. The heat of reaction at standard conditions, $\Delta H_{rxn\,j}^{o}$, will be first calculated and it will require the heat of formation, ΔH_f^{o}, for every component *i* in the reaction at 25°C [5].

$$\Delta H_{rxn\,j}^{\ o} = \sum \nu_i \Delta H_f^{\ o}\Big|_{\text{products}} - \sum \nu_i \Delta H_f^{\ o}\Big|_{\text{reactants}} \qquad 16$$

The heat of reaction at temperature T will then be defined by Equation 17, where the integral term will account for the temperature change.

$$\Delta H_{rxn\ j}(T) = \Delta H_{rxn\ j}{}^{o} + \int_{25^{o}C}^{T} \Delta \hat{C}_{p} dT \qquad 17$$

Due to the isothermal operation of the reactors, the C_p values will be constant along the reactor. Table A-2 contains the appropriate information to calculate the C_p values for each component using the equation at the bottom of the table. The $\Delta \hat{C}_p$ term from Equation 17 could then be written as a constant expression where the Δ term will be defined as the products less the reactants.

$$\Delta C_{p,rxn\ j} = \Delta a + T\Delta b + T^{2}\Delta c + T^{3}\Delta d \qquad 18$$

Using this expression for the $\Delta C_{p,rnx\ j}$, the heat of reaction can be rewritten as follows:

$$\Delta H_{rxn\ j}(T) = \Delta H_{rxn\ j}{}^{o} + \Delta a \Delta T + \frac{1}{2}\Delta b(T_2{}^{2} - T_1{}^{2}) + \frac{1}{3}\Delta c(T_2{}^{3} - T_1{}^{3}) + \frac{1}{4}\Delta d(T_2{}^{4} - T_1{}^{4}) \qquad 19$$

Table A-2. Physical and chemical properties of the species involved in the gas-phase catalyzed reactions.

Species	**MW (kg/ kmol)** [6]	ΔH_f°**(kJ/mol)** [7]	$\boldsymbol{a}$[6]	$\boldsymbol{b}$[6]	$\boldsymbol{c}$[6]	$\boldsymbol{d}$[6]
Propylene	42.081	20410.00	0.886	5.60E-02	-2.77E-05	5.27E-09
O2	32.000	0.00	6.713	-8.79E-07	4.17E-06	-2.54E-09
Steam	18.015	-241830.00	7.701	4.60E-04	2.52E-06	-8.59E-10
Acrolein	56.064	-65000.00	2.859	5.03E-02	-2.56E-05	4.55E-09
Acrylic Acid	72.064	-336900.00	0.416	7.62E-02	-5.62E-05	1.67E-08
CO2	44.010	-393500.00	4.728	1.75E-02	-1.34E-05	4.10E-09
CO	28.010	-110500.00	7.373	-3.07E-03	6.66E-06	-3.04E-09
Acetaldehyde	44.054	-170700.00	1.843	4.35E-02	-2.40E-05	5.69E-09
Acetic Acid	60.052	-433000.00	1.156	6.09E-02	-4.19E-05	1.18E-08
Nitrogen	28.013	0.00	7.44	-3.24E-03	6.40E-06	-2.79E-09

$$\boldsymbol{C_{pi}\left[\frac{J}{mol \cdot K}\right] = \left(a_i + b_i T + c_i T^{2} + d_i T^{3}\right) \times 4.1868}\textbf{, and T is in K}$$

Once the heat of reaction, $\Delta H_{rxn\ j}(T)$, was determined for every reaction, it was multiplied by its respective extent of reaction and the sum of all the terms yield the heat recovery term, $\Delta \dot{H}_{rxn\ j}(T)$,

$$\Delta \dot{H} \qquad \sum_{j} \Delta H_{rxn\ j}(T) \cdot \xi_{j} \qquad 20$$

In the optimization problem, the heat recovery is maximized, but as we have exothermic reactions, the sign of the heat of reaction will be negative. Thus, one will maximize the absolute value of the heat of reaction.

A.4 Multi-Objective Optimization for the Reactor Section

Definition of the Upper and Lower Bounds for the Decision Variables

Inlet Flows to the System

The minimum required production of acrylic acid was set to 60000 metric tons/yr. Assuming a stream factor of 0.962 which corresponds to 50 weeks of operation per year, the production of acrylic acid in a hour basis was 91.40 kmol/h.

$$60000\frac{\text{tons}}{\text{yr}}\cdot\frac{1000\text{kg}}{1\text{ton}}\cdot\frac{\text{kmol}}{72.06\text{kg}}=832\ 593.25\frac{\text{kmol Acryl.A}}{\text{yr}}$$

$$832\ 593.25\frac{\text{kmol Acryl.A}}{\text{yr}}\times\frac{1\text{yr}}{365\text{days}}\cdot\frac{1\text{day}}{24\text{h}}\cdot\frac{351}{365}=91.40\frac{\text{kmol Acryl.A}}{\text{h}}$$

If 100% conversion of acrolein is assumed for the second reactor towards acrylic acid, the minimum amount of acrolein required is 91.40kmol/h. Furthermore, assuming a yield of 90% to acrolein and 100% conversion of propylene in the first reactor, the minimum amount of propylene required would be 101.56 kmol/h.

$$91.40\frac{\text{kmol Ac}}{\text{h}}\cdot\frac{100\%}{90\%}=101.56\frac{\text{kmol Prop}}{\text{h}}$$

To set the bounds for the propylene flow, -10% and +100 of the minimum value was used to set the lower and the upper bounds. Steam is presented in the feed and the amount will normally be about 1 to 15 moles per mole of propylene [8][9]. As for the oxygen, a molar ratio of 1 to 3 moles per mole of propylene is preferred [8]. Using the steam and the oxygen ratios, the lower bound for both was kept at the lowest amount of

propylene whereas for the upper bound, the highest value of the ratio was used. As for the water vapour that is added to the second reactor, it was given a very wide range to allow for further dilution of the mixture going into the second reactor [10].

Operating Conditions of Both Reactors

Bounds on the operating temperatures of both reactors come directly from the experimental conditions under which the rate expressions were determined. 330-430^{o}C for the first reactor and 285-315^{o}C for the second one [11][12][10] .

The pressure for both reactors is reported in literature to be within a range of 1-5 bar [8][13][14][15]. To account for the compression of air to the first reactor, the lower bound was set to 1.05 bar. The lower bound for the pressure of the second reactor was set to 3 bars, since the reactor has a pressure drop and one wanted to avoid having multiple reactors in series. As for the upper bound of both pressures, it was set at 6 bars to allow for flexibility in the optimization problem.

Soft Constraint on the Oxygen Concentration

Flammability limits

Propylene, acrolein, acrylic acid, acetaldehyde and acetic acid (only when at a concentration higher than 80%) are all flammable chemicals [16]. This greatly affects the design of the reactors since there is a safety factor that needs to be accounted for. For this purpose, the lower and upper flammability limits of the chemicals were considered. These limits are relevant for the prevention of explosions inside the reactors since the flammable gas will burn if its concentration falls between these bounds. If the concentration of the flammable gas in air is below the lower flammability limit (LFL), or above the upper flammability limit (UFL), the mixture of the fuel in air cannot burn [17]. However, it is good practice to operate below the LFL given that operating above the UFL can bring the mixture within the limits of combustibility creating an hazardous situation [17]. These limits are usually defined in terms of percent by volume at 25^{o}C and atmospheric pressure.

Effect of Temperature, Pressure and Concentration of Inert Gas

Temperature, pressure and concentration of the inert gas can affect the flammability limits. In terms of temperature, the LFL will decrease with rising temperature since less combustion energy will be required to achieve the limiting flame temperature thus requiring a lower concentration of fuel in air for the mixture to combust [17]. Since the reactors operate way above 25°C, it will be important to account for this change in temperature for the LFL determination. In this sense, Catoire and Naudet [18] developed an empirically based equation to predict the LFL for fuel containing C, H and O atoms at different temperatures (20-400°C) and at atmospheric pressure:

$$LFL(T) = 519.957 \cdot X^{0.70936} \cdot n_C{}^{-0.197} \cdot T^{-0.51536} \qquad 21$$

where

$$X = \frac{1}{1 + 5n_C + \frac{5}{4}n_H - \frac{5}{2}n_O} \qquad 22$$

and where n_C,n_H and n_O are the number of C, H and O atoms in the fuel, respectively.

As for the pressure, its increase will significantly change the UFL while the LFL will be barely affected [17]. Aside from these 2 factors, the concentration of the inert gas will also influence the LFL. In fact, gases with high molar capacities tend to increase the limit as they will act as "thermal ballast" that will quench the flame temperature below the limit flame temperature [17]. Compared to nitrogen, carbon dioxide and water vapour, for example, will be more effective in rendering the mixture non-flammable.

Minimum Oxygen Concentration

Operating outside the flammability limits will always be recommended to avoid any hazardous situation. Nevertheless, as it was mentioned before, the range will tend to change depending on the temperature, pressure and inert gas concentration. For this reason, it will be easier to define a minimum oxygen concentration (MOC), below which flame propagation is not possible [19]. That is to say, if the concentration of oxygen is maintained below the MOC, it will be possible to prevent fires or explosions regardless of the concentration of the fuel. This value is expressed in units of volume percent of

oxygen in the mixture of fuel and air and as the LFL, the MOC will be dependent on temperature, pressure and the inert gas concentration. However, the following equation can be used to estimate the MOC [19]:

$$MOC = z \cdot LEL \qquad 23$$

Where LEL is the lower explosive limit also known as lower flammability limit (LFL) z is the stoichiometric coefficient of oxygen for a combustion reaction of the form of Reaction 1.

$$z = m + \frac{x}{4} - \frac{y}{2} \qquad 24$$

$$C_m H_x O_y + zO_2 \rightarrow mCO_2 + \left(\frac{x}{2}\right) H_2O \qquad R\,1$$

Propylene, which is a flammable gas present in the feed of the propylene oxidation reactor has an LFL and an UFL of 2.4 and 11 vol% in air, respectively, at standard conditions [20]. The MOC in air with nitrogen as the inert gas for propylene as fuel is 11.5 vol%, while it is 14 vol% when the inert gas is carbon dioxide, both at 25°C.[19] Equation 21, was determined for the temperature range of 20 to 400°C, but since the first reactor can operate to as high as 430°C the LFL was calculated at different temperatures (25-430°C) to observe its behaviour over that range.

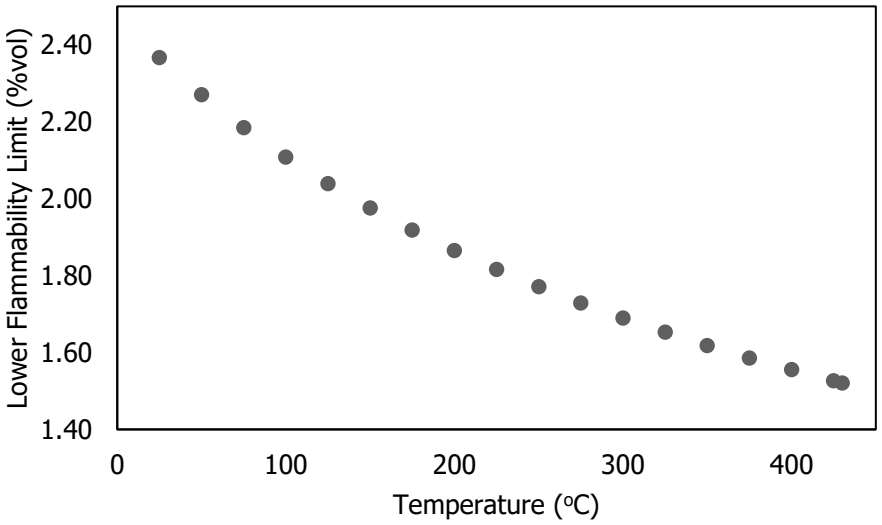

Figure A-2. LFL obtained using Equation 20 [18]***.***

From Figure A-2, it is possible to observe that the LFL follows a very smooth trend over the range of 25-400°C which allows for the extrapolation of the LFL value at 430°C. At 430°C, the LFL for propylene was found to be 1.5 vol%. Using this value, the MOC was calculated using Equation 23, to obtain a value of 6.8 vol% of oxygen in air. Based on this value and knowing that the MOC value will increase when the inert gas is water vapour, one will be able to select the one that was calculated from the correlation directly which will include a safety factor. As for the second reactor, there is no need to put a constraint to avoid flammability. If the limit is respected in the first reactor, it is expected that it will be respected in the second one. Furthermore, water is added into the second reactor which will further dilute the reactor concentration.

In the model, a *sum* term was introduced and initiated to 0. At each iteration, the concentration of the oxygen is evaluated and if larger than 7%vol, the sum term will be modified as follows:

$$sum = sum + (y_{O_2,i} - 0.07)dW$$

where dW is the step size of the catalyst weight.

The *sum* term will then represent the area above the constraint, and as the goal is to keep the oxygen concentration below 7%vol throughout the reactor, this sum will need to be minimized and will be considered as a soft constraint.

References

[1] H. Scott Fogler, *Essentials of Chemical Reaction Engineering*, Prentice Hall, 2011.

[2] R. Turton, R. C. Bailie, W. B. Whiting, and J. A. Shaeiwitz, *Analysis, Synthesis, and Design of Chemical Processes*, 3rd Ed. Upper Saddle River, N.J: Prentice Hall, 2009.

[3] R. Felder And R. Rousseau, *Elementary Principles of Chemical Processes*, 3rd Ed., Vol. 1. John Wiley & Sons, Inc., 2005.

[4] Science & Technology Facilities Council, "Vf04 Hsl Archive," 2011.

[5] J. Smith, H. Van Ness, And M. Abbott, *Introduction to Chemical Engineering Thermodynamics*, 7th Ed. Mcgraw-Hill Education, 2010.

[6] R. C. Reid, J. M. Prausnitz, and T. K. Sherwood, *The Properties of Gases and Liquids*, 3rd Ed. Mcgraw-Hill, Inc, 1977.

[7] U.S Department of Commerce, "Webbook de Chimie Nist," 2017. [Online]. Available: Https://Webbook.Nist.Gov/Chemistry/#Search. [Accessed: 27-Aug-2018].

[8] T. Shiraishi, S. Kishiwada, S. Shimizu, H. Shigern, I. Hiroshi, and N. Yoshihiko, "Catalytic Process for the Preparation of Acrolein," US3970702a, 1976.

[9] T. Noboru Shimizu, I. Isao Yanagisawa, T. Masahiro Takata, and H. Takahisa Sato, "Process for Producing Acrylic Acid from Propylene," 4147885, 1979.

[10] V. C. Malshe and S. B. Chandalia, "Vapour Phase Oxidation of Acrolein to Acrylic Acid on Mixed Oxides as Catalyst," *J. Appl. Chem. Bioiechnol*, Vol. 27, Pp. 575–584, 1977.

[11] H. Redlingshofer, A. Fischer, C. Weckbecker, K. Huthmacher, and G. Emig, "Kinetic Modeling of the Heterogeneously Catalyzed Oxidation of Propene to Acrolein in a Catalytic Wall Reactor," *Ind Eng Chem Res*, Vol. 42, Pp. 5482–5488, 2003.

[12] H. Redlingsho, O. Kro, W. Bo, K. Huthmacher, and G. Emig, "Catalytic Wall Reactor as a Tool for Isothermal Investigations in the Heterogeneously Catalyzed Oxidation of Propene to Acrolein," *Ind Eng Chem Res*, Vol. 41, Pp. 1445–1453, 2002.

[13] O. Hideaki Tsuneki, I. Masanori Nonoguchi, and I. Koji Nishi, "Process for Producing Acrolein, Acrylic Acid and Derivatives Thereof," 9422377b2, 2016.

[14] G. Bub, J. Mosler, D. Maschmeyer, A. Sabbagh, R. Fornika, and M. Peuckert, "Process for The Production of Acrylic Acid," 7294741b2, 13-Nov-2007.

[15] S. A. Setyowati Wibawanta, "Catalytic Partial Oxidation of Propylene for Acrolein

Production," Curtin University, 2011.

[16] N. O. and A. A.-U. Gov, "Cameo Chemicals | Noaa," 2018. [Online]. Available: Https://Cameochemicals.Noaa.Gov/. [Accessed: 29-Aug-2018].

[17] D. Drysdale, *An Introduction to Fire Dynamics*, 3rd Ed. Wiley, 2011.

[18] L. Catoire and V. Naudet, "Estimation of Temperature-Dependent Lower Flammability Limit of Pure Organic Compounds in Air at Atmospheric Pressure," *Process Saf. Prog.*, Vol. 24, No. 2, Pp. 130–137, 2005.

[19] R. H. Perry and D. W. Green, *Perry's Chemical Engineers' Handbook*, 8th Ed. Mcgraw-Hill, 2007.

[20] M. G. Zabetakis, "Flammability Characteristics of Combustible Gases and Vapors," Washington, 1965.

Annex B. Reactor Design Procedure

Detailed design procedure for both reactors

B.1 Propylene Oxidation

Reactions

(1) $\text{Prop}+O_2 \rightarrow \text{Acrolein}+H_2O$

(2) $\text{Acrolein}+0.5O_2 \rightarrow \text{Acrylic A.}$

(3) $\text{Propylene}+2O_2 \rightarrow \text{Acetaldehyde}+H_2O+CO_2$

(4) $\text{Acetaldehyde}+0.5O2 \rightarrow \text{Acetic A.}$

(5) $\text{Propylene}+3O_2 \rightarrow 3CO+3H_2O$

(6) $\text{Propylene}+\frac{9}{2}O_2 \rightarrow 3CO_2+3H_2O$

(7) $\text{Acrylic A.}+3O_2 \rightarrow 3CO_2+2H_2O$

(8) $\text{Acetaldehyde}+\frac{5}{2}O_2 \rightarrow 2CO_2+2H_2O$

(9) $\text{Acetic A.}+2O_2 \rightarrow 2CO_2+2H_2O$

A. Problem Type:

Design: Find Weight of the catalyst

B. Operating Conditions:

- Catalytic Wall Reactor
- Vapour phase
- Isothermal (T = 370°C) and Isobaric (P=1 bar absolute)

C. Design Equations:

$F_{jo} - F_j + \int_W r_j dW = \frac{dN_j}{dt}$, steady state and divide by ΔW

$dF_j + r_j' dW = 0$

$\therefore \frac{dF_j}{dW} = +r_j$

Equation for all the species:

$$\frac{dF_{prop}}{dW}=r_{prop};\frac{dF_{O_2}}{dW}=r_{O_2};\frac{dF_{H_2O}}{dW}=r_{H_2O};\frac{dF_{Acrolein}}{dW}=r_{Acrolein};\frac{dF_{AcrylicA}}{dW}=r_{AcrylicA};$$
$$\frac{dF_{CO_2}}{dW}=r_{CO_2};\frac{dF_{CO}}{dW}=r_{CO};\frac{dF_{Acetaldehyde}}{dW}=r_{Acetaldehyde};\frac{dF_{AceticA}}{dW}=r_{AceticA};\frac{dF_{N_2}}{dW}=0;$$

D. Rate Laws:

(1) $\frac{-r_{1prop}}{1}=\frac{-r_{1O_2}}{1}=\frac{r_{1Acrolein}}{1}=\frac{r_{1H_2O}}{1}$ where $r_{1Acrolein}=k_{1Acrolein}P_{prop}$

(2) $\frac{-r_{2Acrolein}}{1}=\frac{-2r_{2O_2}}{1}=\frac{r_{2AcrylicA}}{1}$ where $r_{2AcrylicA}=k_{2AcrylicA}P_{Acrolein}^{0.86}P_{O_2}^{0.30}$

(3) $\frac{-r_{3prop}}{1}=\frac{-r_{3O_2}}{2}=\frac{r_{3Acetaldehyde}}{1}=\frac{r_{3H_2O}}{1}=\frac{r_{3CO_2}}{1}$ where $r_{3Acetaldehyde}=k_{3Acetaldehyde}P_{O_2}$

(4) $\frac{-r_{4Acetaldehyde}}{1}=\frac{-2r_{4O_2}}{1}=\frac{r_{4AceticA}}{1}$ where $r_{4AceticA}=k_{4AceticA}P_{O_2}^{0.73}$

(5) $\frac{-r_{5prop}}{1}=\frac{-r_{5O_2}}{3}=\frac{r_{5CO}}{3}=\frac{r_{5H_2O}}{3}$ where $r_{5CO}=\frac{k_{5CO}P_{O_2}}{1+K_{H_2O}P_{H_2O}}$

(6) $\frac{-r_{6prop}}{1}=\frac{-2r_{6O_2}}{9}=\frac{r_{6CO_2}}{3}=\frac{r_{6H_2O}}{3}$ where $r_{6CO_2}=\frac{k_{6CO_2}P_{prop}}{1+K_{H_2O}P_{H_2O}}$

(7) $\frac{-r_{7AcrylicA.}}{1}=\frac{-r_{7O_2}}{3}=\frac{r_{7CO_2}}{3}=\frac{r_{7H_2O}}{2}$ where $r_{7CO_2}=\frac{k_{7CO_2}P_{O_2}}{1+K_{H_2O,Acry.A}P_{H_2O}}$

(8) $\frac{-r_{8Acetaldehyde}}{1}=\frac{-2r_{8O_2}}{5}=\frac{r_{8CO_2}}{2}=\frac{r_{8H_2O}}{2}$ where $r_{8CO_2}=k_{8CO_2}P_{Acetaldehyde}$

(9) $\frac{-r_{9AceticA.}}{1}=\frac{-r_{9O_2}}{2}=\frac{r_{9CO_2}}{2}=\frac{r_{9H_2O}}{2}$ where $r_{9CO_2}=k_{9CO_2}P_{AceticA.}$

Net Rates of Formation:

$$r_{prop}=r_{1prop}+r_{3prop}+r_{5prop}+r_{6prop}$$
$$r_{O_2}=r_{1O_2}+r_{2O_2}+r_{3O_2}+r_{4O_2}+r_{5O_2}+r_{6O_2}+r_{7O_2}+r_{8O_2}+r_{9O_2}$$
$$r_{H_2O}=r_{1H_2O}+r_{3H_2O}+r_{5H_2O}+r_{6H_2O}+r_{7H_2O}+r_{8H_2O}+r_{9H_2O}$$
$$r_{Acrolein}=r_{1Acrolein}+r_{2Acrolein}$$
$$r_{AcrylicA.}=r_{2AcrylicA.}+r_{7AcrylicA.}$$
$$r_{CO_2}=r_{3CO_2}+r_{6CO_2}+r_{7CO_2}+r_{8CO_2}+r_{9CO_2}$$
$$r_{CO}=r_{5CO}$$
$$r_{Acetaldehyde}=r_{3Acetaldehyde}+r_{4Acetaldehyde}+r_{8Acetaldehyde}$$
$$r_{AceticA.}=r_{4AceticA.}+r_{9AceticA.}$$

Replacing the rate laws into the design equations:

$$\frac{dF_{prop}}{dW} = -r_{1Acrolein} - r_{3Acetaldehyde} - \frac{1}{3}r_{5CO} - \frac{1}{3}r_{6CO_2}$$

$$\frac{dF_{O_2}}{dW} = -r_{1Acrolein} - \frac{1}{2}r_{2AcrylicA.} - 2r_{3Acetaldehyde} - \frac{1}{2}r_{4AceticA.} - r_{5CO} - \frac{3}{2}r_{6CO_2} - r_{7CO_2} - \frac{5}{4}r_{8CO_2} - r_{9CO_2}$$

$$\frac{dF_{H_2O}}{dW} = r_{1Acrolein} + r_{3Acetaldehyde} + r_{5CO} + r_{6CO_2} + \frac{2}{3}r_{7CO_2} + r_{8CO_2} + r_{9CO_2}$$

$$\frac{dF_{Acrolein}}{dW} = r_{1Acrolein} - r_{2AcrylicA.}$$

$$\frac{dF_{AcrylicA.}}{dW} = r_{2AcrylicA.} - \frac{1}{3}r_{7CO_2}$$

$$\frac{dF_{CO_2}}{dW} = r_{3Acetaldehyde} + r_{6CO_2} + r_{7CO_2} + r_{8CO_2} + r_{9CO_2}$$

$$\frac{dF_{CO}}{dW} = r_{5CO}$$

$$\frac{dF_{Acetaldehyde}}{dW} = r_{3Acetaldehyde} - r_{4AceticA.} - \frac{1}{2}r_{8CO_2}$$

$$\frac{dF_{AceticA.}}{dW} = r_{4AceticA.} - \frac{1}{2}r_{9CO_2}$$

E. Concentration Expressions: The ones in the rate expressions

$$C_j = C_{To}\frac{F_j}{F_T} \text{ and } P_j = C_{To}\frac{F_j}{F_T}RT$$

$$P_{prop} = C_{To}\frac{F_{prop}}{F_T}RT;\ P_{O_2} = C_{To}\frac{F_{O_2}}{F_T}RT;\ P_{H_2O} = C_{To}\frac{F_{H_2O}}{F_T}RT;\ P_{Acrolein} = C_{To}\frac{F_{Acrolein}}{F_T}RT;$$

$$P_{Acetaldehyde} = C_{To}\frac{F_{Acetaldehyde}}{F_T}RT;\ P_{AceticA.} = C_{To}\frac{F_{AceticA.}}{F_T}RT$$

F. Parameter Evaluation:

- Units:

Molar flowrate: $F_j = \frac{kmol}{h}$

Rate law: $\frac{dF_j}{dW} = r_i = \frac{kmol}{kgcat * s} \times \frac{3600s}{1h} = \frac{kmol}{kgcat * h}$

Partial pressure: $P_j = \frac{kmol}{m^3} \times \frac{kmol}{h} \times \frac{m^3 * Pa}{mol * K} \times K \times \frac{h}{kmol} \times \frac{1000mol}{1kmol} \times \frac{1bar}{100000Pa} = bar$

B.2 Acrolein Oxidation

Reactions

(1) $\text{Acrolein} + \frac{1}{2}\text{O}_2 \rightarrow \text{AcrylicA.}$

(2) $\text{Acrolein} + \frac{7}{2}\text{O}_2 \rightarrow 3\text{CO}_2 + 2\text{H}_2\text{O}$

A. Problem Type:

Design: Find Weight of the catalyst

B. Operating Conditions:

- PBR
- Vapour phase-ΔP from Ergun equation
- Isothermal (T=285^{o}C)
- Kinetics were found for a feed containing only air, acrolein and water. The other components present in the feed will be accounted in the nitrogen term (inerts).

C. Design Equations:

Equation for all the species:

$$\frac{dF_{Acrolein}}{dW} = r_{Acrolein}; \frac{dF_{O_2}}{dW} = r_{O_2}; \frac{dF_{H_2O}}{dW} = r_{H_2O}; \frac{dF_{AcrylicA}}{dW} = r_{AcrylicA}; \frac{dF_{CO_2}}{dW} = r_{CO_2}; \frac{dF_{N_2}}{dW} = 0;$$

D. Rate Laws:

(1) $\frac{-r_{1Acrolein}}{1} = \frac{-2r_{1O_2}}{1} = \frac{r_{1AcrylicA.}}{1}$ where $r_{1AcrylicA.} = \frac{k_{1AcrylicA.}C_{Acrolein}}{1 + K_2 C_{AcrylicA.}}$

(2) $\frac{-r_{2Acrolein}}{1} = \frac{-2r_{2O_2}}{7} = \frac{r_{2CO_2}}{3} = \frac{r_{2H_2O}}{3}$ where $r_{2CO_2} = \frac{k_{2CO_2}C_{Acrolein}}{1 + K_4 C_{AcrylicA.}}$

Net Rates of Formation:

$$r_{Acrolein} = r_{1Acrolein} + r_{2Acrolein}$$

$$r_{O_2} = r_{1O_2} + r_{2O_2}$$
$$r_{H_2O} = r_{2H_2O}$$
$$r_{AcrylicA.} = r_{1AcrylicA.}$$
$$r_{CO_2} = r_{2CO_2}$$

Replacing the rate laws into the design equations:

$$\frac{dF_{Acrolein}}{dW} = -r_{1AcrylicA.} - \frac{1}{3} r_{2CO_2}$$

$$\frac{dF_{O_2}}{dW} = -\frac{1}{2} r_{1AcrylicA.} - \frac{7}{6} r_{2CO_2}$$

$$\frac{dF_{H_2O}}{dW} = \frac{2}{3} r_{2CO_2}$$

$$\frac{dF_{AcrylicA.}}{dW} = r_{1AcrylicA.}$$

$$\frac{dF_{CO_2}}{dW} = r_{2CO_2}$$

E. Concentration Expressions: The ones in the rate expressions

$$C_j = C_{To} \frac{F_j}{F_T} \frac{P}{P_o}$$

$$C_{Acrolein} = C_{To} \frac{F_{Acrolein}}{F_T} \frac{P}{P_o} \; ; \; C_{AcrylicA.} = C_{To} \frac{F_{AcrylicA.}}{F_T} \frac{P}{P_o}$$

F. Ergun Equation:

$$\frac{dy}{dW} = -\frac{\alpha}{2y} \frac{F_T}{F_{To}}$$

Where

$$y = \frac{P}{P_o}$$

$$\alpha = \frac{2\beta_o}{A_c \rho_c (1-\phi) P_o}$$

$$\beta_o = \frac{G(1-\phi)}{\rho_o g_c D_p \phi^3}\left[\frac{150(1-\phi)\mu}{D_p} + 1.75G\right]$$

G. Parameter Evaluation:

- Units:

Molar flowrate: $F_j = \frac{kmol}{h}$

Rate law: $\frac{dF_j}{dW} = r_i = \frac{kmol}{m^3 cat * s} \times \frac{3600s}{1h} \times \frac{1}{\rho_c}$

$$= \frac{kmol}{m^3 cat * s} \times \frac{3600s}{1h} \times \frac{m^3 cat}{kgcat}$$

$$= \frac{kmol}{kgcat * h}$$

Pressure drop:

$$\beta_o = \frac{\frac{kg}{m^2 s}}{\frac{kg}{m^3} \times m}\left[\frac{Pa \times s}{m} + \frac{kg}{m^2 s}\right]$$

$$= \frac{\frac{kg}{m^2 s}}{\frac{kg}{m^2}}\left[\frac{\frac{kg \times s}{m^2 s^2} \times s}{m} + \frac{kg}{m^2 s}\right]$$

$$= \frac{Pa}{m} \times \frac{1\ kPa}{1000\ Pa}$$

$$= \frac{kPa}{m}$$

$$\alpha = \frac{\frac{kPa}{m}}{m^2 \frac{kg}{m^3} kPa} = \frac{1}{kg}$$

$$\frac{dy}{dW} = -\frac{\frac{1}{kg}}{2y} \frac{\frac{kmol}{h}}{\frac{kmol}{h}} = -\frac{1}{kg}$$