Tradespace as a Decision Making Tool in Bioprocess Design

by

Yu-wen Tang

B.S. Molecular Biotechnology Engineering
Universidad de Chile, 2011

SUBMITTED TO THE SYSTEM DESIGN AND MANAGEMENT PROGRAM
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN ENGINEERING AND MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2016

Signature redacted

Signature of Author_____
Yu-wen Tang
System Design and Management Program
May 20th, 2016

Signature redacted

Certified by_____
Andrey Zarur
Senior Lecturer, Martin Trust Center for MIT Entrepreneurship, Sloan School of Management
Thesis Supervisor

Signature redacted

Accepted by_____
Patrick Hale
Director
System Design and Management Program

(Page intentionally left in blank)

Tradespace as a Decision Making Tool in Bioprocess Design
by

Yu-wen Tang

Submitted to the System Design and Management Program on
May 20th, 2016 in Partial Fulfillment of the Requirements for
the Degree of Master of Science in Engineering and
Management

ABSTRACT

The field of systems engineering upholds that fundamental engineering principles exist and are applicable across different domains and contexts. In this thesis, a state-of-the art decision and design evaluation method developed for aerospace systems, Multi-Attribute Tradespace Exploration (MATE) is complemented with Design of Experiments (DoE) and applied for the first time to a bioprocess design problem. The implementation of DoE was necessary due to the high complexity of bioprocess systems, where a design variable (or a reasonably small number of design variables) cannot be easily identified to explain a given attribute of the product or process. DoE not only allows the identification of design variables that most influence a given attribute, but also allows the development of Single-Utility-Functions facilitating the incorporation of the Multi-Utility component of the MATE method.

The proposed new MATE-DoE method was implemented in two case studies to assess its applicability; namely bio-production of DHA and bio-production of a lipase enzyme. Based on published DoE experimental results, utility functions and cost estimations were carried out to develop a Tradespace. The resulting Tradespace demonstrates: (a) the possible implementation of the proposed method, (b) that the use of Tradespace complements the traditional bioprocess development practice by allowing decision makers to choose an architecture that optimizes for more than one objective (multi-objective), (c) that the proposed method takes into consideration the complex decision making process of customers (multi-attribute), and (d) that simultaneous comparison analysis to competitors and market standards are possible using the method.

While the method was proven to be applicable, it is relatively complex and the number of experiments and market data required might prevent its broad implementation. Also, potential errors and misleading results might result from inaccurate input data. Special attention and effort need to be put in accurate Single-Utility Function (SUF) weight designation to avoid this problem. The importance of assessing the complete bioprocess, as opposed to individual unit operations, is highlighted. Finally, further studies to develop "rules of thumb" in order to simplify the proposed MATE-DoE method is suggested.

Thesis Supervisor: Andrey Zarur
Title: Senior Lecturer, Martin Trust Center for MIT Entrepreneurship, Sloan School of Management

ACKNOWLEDGEMENTS

I would like to thank my advisor, Andrey Zarur for his support, patience and understanding throughout the development of this thesis. For allowing and encouraging me to try something new and pursue my intellectual curiosity. Also, special thank you to Karthikeyan D. Ramachandriya for keeping me on track, and for providing his experience and guidance.

I am grateful to SDM and MIT community for this life changing experience. The SDM program provides students with a very valuable peer group, from different academic background, work histories and interests. The interaction, learnings and support from this group has been invaluable and I am forever thankful for their friendship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. PROBLEM AND MOTIVATION

Industrial biotechnology encompasses the application of biotechnology-based tools to traditional industrial processes ("bioprocessing") and the manufacturing of bio-based products (such as fuels, chemicals and plastics) from renewable feedstock (Erickson, Nelson, & Winters, 2012). The development of technologies and disciplines, such as second generation genome sequencing, synthetic biology and fermentation process engineering has opened the door for the development of new biochemical platforms. Process optimization is crucial for platforms based on fermentation technologies. This is especially true for biotech startups focused on the development of more sustainable, environmentally-friendly and cost-efficient commodity chemicals. However, traditionally bioprocess optimization is highly focused on increasing productivity, while cost analysis studies are often times performed later in the PDP (Product/Process Development Process). In other words, it is a lineal process were production parameters are first defined and then a cost estimation for the established parameter is performed. This lineal process misses the opportunity of operating cost optimization. For example increasing the agitation rate (RPM) in a bioreactor can increase aeration, resulting in increase of yield (biomass), but it also increases power consumption and therefore cost (Gill, Appleton, Baganz, & Lye, 2008). Thus, the question remains, what is the choice of conditions that optimizes both, yield and cost? Furthermore, how can other attributes such as purity and stability be accounted for at the same time yield is maximize and cost minimized? The following thesis proposes the use of Tradespace Analysis study as a dynamic and visual tool that can support the decision of what the best production parameters are that optimize overall performance and process cost simultaneously. Furthermore, the use of such tool can help close the gap between engineering and management by articulating clearly the magnitude of realized savings in the optimization of a given process and justify investment in process development.

The present thesis explores the applicability of Tradespace Analysis by developing a conceptual Tradespace process for the production of biocompounds. In order to partially illustrate how the suggested Tradespace process could be applied, two case studies were analyzed, namely the production of 1) a bacterial lipase enzyme and 2) an algal DHA. In the following subsections the context in which this two biocompound are used, their economic importance, as well as a general overview of bioprocess, fermentation and PDP in biochemical engineering is presented.

Finally, the applicability of Tradespace Analysis in assisting the PDP process and how it differs from other bioprocess modelling tools is discussed.

1.2. INDUSTRIAL BIOTECHNOLOGY

Industrial biotechnology (also known as white biotechnology or green chemistry) refers to the use of living systems, organisms or components of cells, such as enzymes, to develop and make products in sectors such as chemical, food and feed, detergent, pulp and paper, textile and bioenergy (such as biofuel and biogas). It is often referred to as the third wave of biotechnology (The Economist , 2009). The first wave refers to biotech products in medicine; second wave refers to agricultural biotechnology. Industrial biotechnology is one of the most promising approaches to pollution prevention, resource conservation and lowering greenhouse gas emissions. By using renewable raw materials bioprocesses are cleaner, more sustainable and contribute to moving away from petrochemical-based economy to a biobased economy. "It offers businesses a way to reduce cost and create new markets while protecting the environment. Also, since many of its products do not require the lengthy review times that drug products must undergo, it's quicker, easier pathway to the market. Today, new industrial processes can be taken from the lab to the commercial application in two to five years, compared to up to a decade for drugs" (Simpson, 2005).

1.3. INDUSTRIAL BIOTECHNOLOGY IN FOOD AND FEED

The use of biological processes in food production existed long before the discovery of microorganisms. It was perhaps the very first application of biotehcnology. The oldest biological process in food production is the conversion of sugar to alcohol by yeasts to produce beer– a fermentation process carried out as early as 7000 BC on small and individual scale by Chinese villagers (McGovern, Zhang, & J.G, 2004). However, the process of fermentation was not fully understood until it was described much later, by Louis Pasteur in 1857, who concluded that fermentation was a living process of yeasts (Demain, 2010).

Humans have leveraged fermentation processes and used microorganisms such as yeast and bacteria for the production and preparation of foods for thousands of years. Amongst the traditional products produced through fermentation are bread, wine, beer, yogurt, cheese, sausage, soy sauce, vinegar, amongst others. With the advent of molecular biology, microorganisms are today being

genetically engineered and used as cell factories to produce a great variety of metabolites and enzymes that are used as food, drink or food additives (Table 1) (Rahman, 2016).

Table 1 – Food and food ingredients produced by microorganisms in fermentation industries.
Source: Adapted from *(Rahman, 2016)*

Biochemical group	Product type	Product	Organism
Biomass	Starter culture	Cultured buttermilk, cultured sour cream	*Lactococcus lactis* or *Streptococcus cremoris* and *Leuconostoc citrovorum* or *Leuconostoc dextranicum* (mixed)
		Bulgarian milk	*Lactobacillus bulgaricus*
		Acidophilus milk	*Lactobacillus acidophilus*
		Yogurt	*Streptococcus thermophiles* *Lactobacillus bulgaricus*
		Baker's yeast	*Saccharomyces cervisiae*
	Probiotics	Fermented milk, yogurt, cheese, dried powder, captuse	*Lactobacillus casei* Shirota *Lactobacillus johnsonii* *Lactobacillus casei* *Bifidobacterium animalis* *Lactobacillus acidophilus* NCFM *Streptococcus thermophilus,* *Streptococcus thermophilus,* *Enterococcus faecium*
Primary metabolite	Alcoholic beverage	Beer, wine	*Saccharomyces cervisiae*
	Preservative/ acidulants	Lactic acid	*Lactococcus lactis*
Intermediate metabolite	Nutritional supplements/ neutraceutical	L-tryptophan, L-phenylalanine, L-tyrosine, L-threonine, L-isoleucine, L-histidine, vitamins	*Corynebacterium glutamicum* Lactic acid bacteria *Bacillus subtillis*
	Flavor enhancer/ aroma compound	Isoprenoids, diacetyl, acetaldehyde	*Saccharomyces cervisieae* Lactic acid bacteria
	Sweetener	Xylitol, L-alanine, Mannitol, Sorbitol	*Saccharomyces cervisieae* Lactic acid bacteria
	Preservatives/ acidulants	Citric acid, acetic acid lactic acid, succinic acid, pyruvate	Lactic acid bacteria *Escherichia coli*
Secondary metabolites (non-antibiotic)	Functional food/ neutraceuticals	Resveratol (flavonoid)	*Saccharomyces cervisiae* *Escherichia coli*
	Food coloration	Carotenoids	*Saccharomyces cervisiae* *Escherichia coli*
	Prebiotics	Exopolysaccharides	Lactic acid bacteria
Secondary metabolite (antibiotic)	Preservatives	Bacteriocins	*Bacillus subtillis*
Enzymes	Lipases, Amylase, Galactosidase Etc.	Used in several food production processes. See table 2.	

From the list above, production of enzymes for food processing are of special interest, as they are used to improve a great variety of food manufacturing processes (Table 2). They can improve texture, appearance, nutritional value and may generate desirable flavors and aromas. Because they are used in such a wide variety of food and beverage production processes, the food enzymes market is expected to reach US$ 2.7 billion by 2020, with a CAGR of 8.1% from 2015-2020 (Mordor Intelligence, 2015).

Table 2 – Enzymes used in food and beverage industry and their application.

Source: (Amore & Faraco, 2016)

Industry	Enzyme	Application
Starch	Amylase	Starch liquefaction and saccharification
	Amyloglucosidase	Saccharification
	Pullulanase	Saccharification
	Glucose isomerase	Glucose to fructose conversion
	Cyclodextrin-glycosyltransferase	Cyclodextrin production
	Xylanase	Viscosity reduction (starch)
Food (including dairy)	Protease	Milk clotting, infant formulas (low allergenic), flavour
	Lipase	Cheese flavour
	Lactase	Lactose removal (milk)
	Pectin methyl esterase	Firming fruit-based products
	Pectinase	Fruit-based products
	Transglutaminase	Modify visco-elastic properties
Baking	Amylase	Bread softness and volume, flour adjustment
	Xylanase	Dough conditioning
	Lipase	Dough stability and conditioning (*in situ* emulsifier)
	Phospholipase	Dough stability and conditioning (*in situ* emulsifier)
	Glucose oxidase	Dough strengthening
	Lipoxygenase	Dough strengthening, bread whitening
	Protease	Biscuits, cookies
Beverage	Pectinase	De-pectinization, mashing
	Amylase	Juice treatment, low-calorie beer
	ß-glucanase	Mashing
	Acetolactate decarboxylase	Maturation (beer)
	Laccase	Clarification (juice), flavour (beer), cork stop

Food enzymes are typically derived from animal, vegetable and microbial sources. However, both animal and vegetable enzymes present difficulty in the extraction process and other products of cellular metabolism can interfere with the enzymatic activity. Also, vegetable enzyme production depends on external factors such as climate, soil and seed. Furthermore, animal enzymes are limited by ethical parameters, shortage and concerns related to health and origin of the animals and organs (Vermelho, Cardoso, Pires Nascimento, Pinheiro, & Rodriguez, 2016). In contrast, microbial enzymes production proceeds independently of external factors, has a simple extraction process and does not have shortage or animal health concerns. The main sources of microbial enzymes are bacteria, fungi and yeast. These are grown under controlled conditions in bioreactors. At the end of the process, the broth in the bioreactor contains enzymes, nutrients and

the corresponding microorganism, from which the desired enzyme is extracted and purified (for more details please refer to section 1.4). Moreover, with the advance of synthetic biology, new enzymes, with new functions and improved production are being introduced into the food industry. However, despite the many advantages of microbial enzymes, their overall production cost is high relative to those of vegetal and animal enzymes origin. This can significantly limit their introduction into the market.

Yet another group of microorganisms used as cell factories, as well as raw materials for food and feed products, are microalgae. Even though the use of microalgae by humans date back 2000 years to the Chinese, who used *Nostoc* to survive during famine (Spolaore, Joannis-Cassan, Duran, & Isambert, 2006), it was not until the 1940s that microalgae became more and more important as live feeds in aquaculture (Hallmann, 2007). And while macroalgae (seaweed), have an old tradition in the use of biomass for the production of phycocolloids like aga-agar, alginates or carrageenan (Pulz & Gross, 2004), the use of microalgae in biotechnology is significantly more recent. As a matter of fact, from virtually none in 1990, the total number of publications on algal biotechnology leapt to 153 by June 2011; of these, 103 were on microalgal biofuel (Darvasula, Darvasula V., & Rao, 2013). The market size of products from microalgae was estimated by Pultz and Gross in 2004 to have a retail value of US$ 5-6.5 billion (Table 3). From which the biggest segment were biomass for health food and production of docosahexanoic acid (DHA), US$1.25-2.5 billion and US$1.5 billion, respectively.

Table 3 –Market estimation for microalgal products

Source: Adapted from (Pulz & Gross, 2004)

Product group	Product	Retail value in 2004 (USD x10^6)
Biomass	Health food	1,250-2,500
	Functional food	800
	Feed additive	30
	Aquaculture	700
Coloring substances	Astaxanthin	<150
	Phycocyanin	>10
Antioxidants	B-Carotene	>280
	Antioxidant extract	100-150
Fatty acids	ARA	20
	DHA	1,500
	Poly-unsaturated fatty acid extracts	10

Microalgae (as well as other types of algae and cyanobacteria) are mainly cultivated by two approaches: a) Open ponds or tanks and b) closed bioreactors (Figure1). The first relies on the use of solar energy to produce biomass; thus a photosynthetic process. The second requires the

input of nutrients (carbon and nitrogen sources), heat and in some cases oxygen; thus a fermentation process. Bioreactors provide better control of growth parameters, prevent contamination and allow higher volumetric productivities, at the expense of higher cost and energy requirements than open pond (Harun, Singh, Forde, & Danquah, 2010). Since both systems have benefits and limitations, the choice will depend on the final product, targeted market, biomass productivity and metabolic requirements of the specific microalgae strain used.

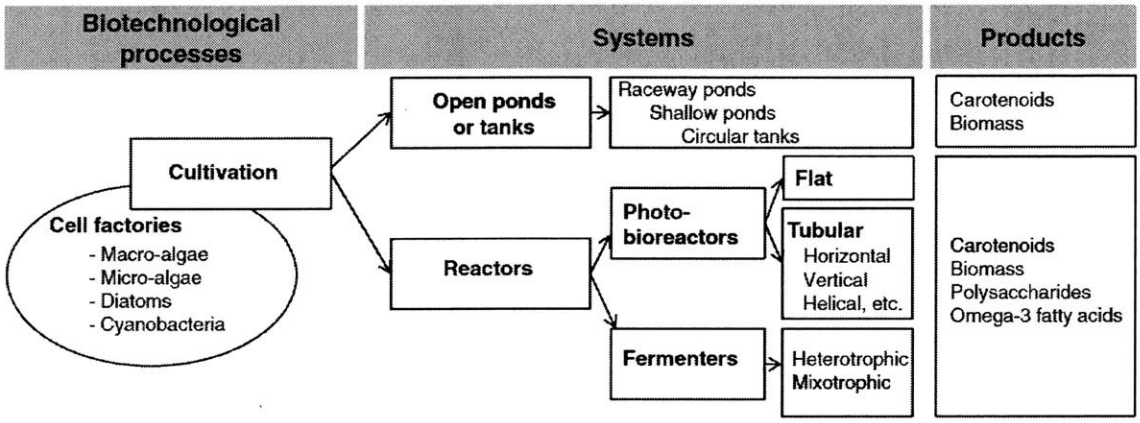


Figure 1 – Marine microorganism as cell factories.

Source: Adapted from (Freitas, Rodrigues, Rocha-Santos, Gomez, & Duarte, 2012)

"There are two main categories of food market products obtained from microalgae. The first category is dried algae (in particular the micro-algae species Chlorella and Spirulina) with high nutrient content, especially of vitamin B12, C and D2. These micro-algal products can be directly sold as dietary supplements and have the potential to be used in bulk commodities as sources of protein and carbohydrates. The second type is specialty products isolated and extracted from the micro-algae that can be added to food and feed to improve their nutritional value. These high-value compounds are pigments (e.g. astaxanthin), anti-oxidants (e.g. β-carotene), protein (e.g. phycocyanin) and fatty acids (e.g. docosahexaenoic acid –DHA and eicosapentaenoic acid –EPA)" (Vigani, et al., 2015).

In this thesis, two food products produced through fermentation were used as case studies for the analysis of Tradespace as a decision making tool in bioprocesses design. Both products belong to the fat and oil segment. However, the first one is used as part of the process to obtain higher value fats, namely lipases, and the second is a high value fatty acid (DHA). In the following

subsections, both lipase and DHA importance, market size and application in food industry are introduced.

1.3.1. LIPASE

Lipases (triacylglycerol acylhydrolases EC: 3.1.1.3) are some of the most useful enzymes for food processing (Jaeger & Eggert, 2002). Lipases catalyze the hydrolysis (cleavage of chemical bond by addition of water) and synthesis of lipids (Figure 2). Lipases are ubiquitous enzymes found in animals, plants, fungi and bacteria. However, microbial lipases are commercially significant as they are of low production cost, greater stability and wider availability than plant and animal lipases (Aravindan, Anbumathi, & Viruthagiri, 2007). Also, microbial lipases are of great biotechnological interest because these enzymes are: (1) stable in organic solvents, (2) do not require cofactors, (3) have great substrate specificity, (4) act over wide range of pH and temperature and (5) have a high enantioselectivity[1] (Andualema & Gessesse, 2012)

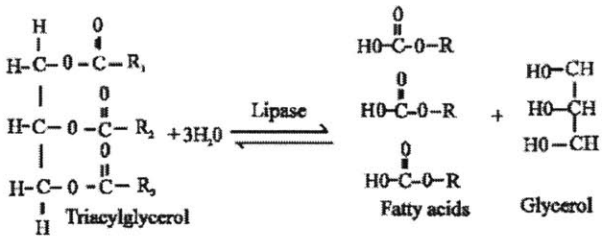


Figure 2 – Lipase catalyzed reaction.
A triglyceride can be hydrolyzed to form glycerol and fatty acids, or the reverse (synthesis) reaction can combine glycerol and fatty acids to form the triglyceride. Source: (Andualema & Gessesse, 2012)

In terms of market size, "the market for lipase is projected to reach $590.5 Million by 2020, at a CAGR of 6.5% between 2015 and 2020. The global lipase market is expected to grow in the near future, owing to factors such as increasing health awareness among people across the globe, changing dietary habits and growing technological advances in the food and beverage industry" (Markets and Markets, 2015). Lipases are considered to be the third largest commercial enzyme

[1] According to Merriam Webster Medical Dictionary:
Enantioselectivity: the degree to which one enantiomer of a chiral product is preferentially produced in a chemical reaction.
Enantiomer: either of a pair of chemical compound whose molecular structure have a mirror-image relationship to each other – called also optical antipode
Chiral:a) having a structure that is nonsuperimposable on its mirror image <chiral molecule> b)relating to or composed of chiral molecules

group based on total sales volume, after proteases and carbohydrases (Andualema & Gessesse, 2012).

Within the food and beverage industry, lipases can be used as additives or as biocatalysts to manufacture food ingredients. As additives, lipases can hydrolyze fats into short-chained, esters fatty acids and alcohols, which are known flavor and fragrance compounds. This is a common practice in the dairy industry, where lipases are used to hydrolyze milk fat to enhance cheese flavor, accelerate cheese ripening, in the manufacturing of cheese-like products and the lipolysis of butterfat. The addition of lipases in these processes allows the release of short chain fatty acids (primarily C4 and C6) leading to the development of a sharp, tangy flavor and the digestion of medium chain (C12, C14) fatty acids, which tend to confer a soapy taste to the product (Ferreira-Dias, Sandoval, Plou, & Valero, 2013).

As biocatalysts, lipases are important in the lipid industry because they can be exploited for the retailoring of vegetable and animal oils. "Fats and oils are important constituents of food. The nutritional and sensory value and the physical properties of a triglyceride are greatly influenced by factors such as the position of the fatty acid in the glycerol backbone, the chain length of the fatty acid, and the degree of unsaturation. Lipases allow us to modify the properties of lipids by altering the location of fatty acid chains in the glyceride and replacing one or more of the fatty acids with new ones. This way, a relatively inexpensive and less desirable lipid can be modified to a higher value fat" (Sharma, Christi, & Banerjee, 2001). "For example, cocoa butter fat required for chocolate production is often in short supply and the price fluctuates widely. However, lipase-catalyzed trans-esterification of cheaper oils can be used, for example to produce cocoa butter from palm mid-fraction" (Andualema & Gessesse, 2012). In this way, cheap oils could be upgraded to synthesize nutritionally important structured triacylglycerol (like cocoa butter substitutes), low calorie triacylglycerol, human milk fat substitutes and oils enriched with specific fatty acids such as oleic, stearidonic, gamma-linoleic (GLA), conjugated linoleic (CLA) or omega-3 polysaturated (ω3 PUFA) fatty acid (Ferreira-Dias, Sandoval, Plou, & Valero, 2013).

The first case study in this thesis is a hyperthermostable lipase from Bukholderia cepacia (Rathi, Saxena, & Gupta, 2001). The enzyme is already known to have a broad temperature range of 25-100 °C and exhibits the novel phenomenon of thermal activation. High thermostability is important, because in high temperatures substrate solubility increases, at the same time as the viscosity decreases and thereby avoids environmental contamination (Andualema & Gessesse,

2012). Also, the use of thermostable lipases is very important when these reactions occur in solvent-free media, where at least one of the substrates has a high melting point (m.p.), such as palm stearin (m.p. = 47-54°C). To carry out these reactions at near-room temperature, an organic solvent to dissolve the solid fats is needed. This will increase the complexity of the system, as well as the costs related with solvent and downstream processing. In the last decade, these facts, together with the search for green processes, have drawn special attention to the search for lipases produced by thermophilic microorganisms (Ferreira-Dias, Sandoval, Plou, & Valero, 2013)

1.3.2. DHA

Docosahexaenoic acid (DHA) is an omega-3 fatty acid that is a primary structural component of the human brain, cerebral cortex, skin, sperm, testicles and retina. It can be synthesized from alpha-linolenic acid or obtained directly from maternal milk (breast milk), fish oil or algae oil (Guesnet & Alessandri, 2011). Traditionally DHA was produced from fish oil, however, fish oil DHA has limited applications as an additive because of its smell, unfavorable fishy flavor and weak oxidative stability. Moreover, the environmental implications due to overfishing has promoted the effort to put more restrictive laws and fishing quotas in place, limiting the amount of fish oil available for DHA (Iacurci, 2014). Also, fish oil is unsuitable for neonate formula because of the presence of eicosapentanoic acid (EPA), which acts antagonistically with arachidonic acid (ARA) (De Swaaf, De Rijk, Eggink, & Sijtsma, 1999). Thus, algae is currently the major alternative source for production of omega-3, especially DHA. According to Frost and Sullivan (2014), algae DHA market is valued at US$329 million in 2012 and is expected to grow at a 12.3% CAGR to US$1,175 million in 2023. The demand in 2012 was estimated to be 4,614 metric tons, thus an average price of US$ 54.91 per kg of DHA. DHA is predominantly used as additive in infant formula (Figure 3), accounting for 50% (US$165.23 million) of microalgae DHA revenue and 48.9% of its shipment (Figure 3).

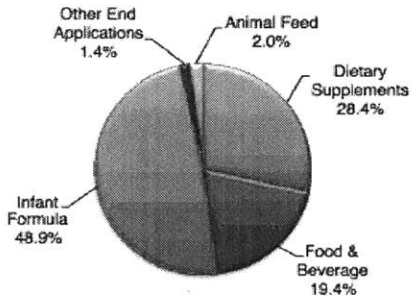
Microalgae DHA Oils Segment—End Application Dynamics

The infant formula application is only expected to grow at a CAGR of 4.3% because it has already reached saturation in most major markets.

Microalgae DHA Oils Segment: Volume Shipment, Revenue and CAGRs by End Application, Global, 2012–2023

End Application	2012 Revenue (€ Million)	Revenue CAGR (2012–2023) (%)	2012 Volume Shipment (Metric Tons)	Volume Shipment CAGR (2012–2023) (%)
Dietary Supplements	58.2	23.4	1,309	22.7
Food & Beverage	60.4	11.3	894	16.5
Infant Formula	127.1	(0.7)	2,254	4.3
Animal Feed	2.1	18.7	92	20.7
Other End Applications	5.2	5.3	65	11.2
Total	**253.0**	**12.3**	**4,614**	**15.0**

Microalgae DHA Oils Segment: Percent Volume Shipment by End Application, Global, 2012



Other end applications include personal care, pharmaceuticals, and cosmetics
Note: All figures are rounded. The base year is 2012. Source: Frost & Sullivan

NE34-88 FROST & SULLIVAN 32

Figure 3 – Microalgae DHA Oil market segment
Conversion rate 1 euro = US$1.3. Source: (Frost and Sullivan, 2014)

Main species of microalgae used for DHA commercial production are *Crypthecodinium cohnii* and *Schizochytrium* (Khosravi-Darani, Koohy-Kamaly, Nikoopour, & Zeinab Asadi, 2016). DSM, became the market leader in algal production of DHA after acquiring Martek Biosciences Corporation for US$1.087 million in 2010 (Heerlen, 2010). The estimated production capacity of algae DHA by DSM was 5,000 metric tons in 2012.

The second case study in this thesis is DHA production of an algae strain, namely *Schizochytrium limacinum* OUC88. This is a heterotrophic microbe (non-photosynthetic), with high content of DHA (Table 4)

Table 4 –Fatty acid composition (% of total fatty acid) of *Schizochytrium limacinum* OUC88
Source: (Song X. , Zhang, Guo, Zhu, & Kuang, 2007)

Fatty acids	Content (%)
14:0	8.34
15:0	1.65
16:0	37.9
17:0	0.85
18:0	1.90
18:2 *n*–6	0.23
18:3 *n*–3	0.49
18:3 *n*–6	0.27
20:0	0.41
21:0	0.27
20:3 *n*–6	0.35
20:4 *n*–6	0.36
22:0	0.38
20:5 *n*–3 (EPA)	0.76
22:5 *n*–6 (DPA)	8.22
22:6 *n*–3 (DHA)	37.5

All data are means of three replicates.

1.4. BIOPROCESSES

Bioprocess can be referred to as a method or operation that uses a living system or their components to produce commercially useful products. The fundamental operational element in a bioprocess is the enzyme, while the scope of the bioprocess ranges from reactions with single enzymes, mixture of enzymes, singles cells to even animal and plant systems (Fig 4). Both case studies researched in this thesis refers to cell cultivations, specifically a bacterial lipase and an algal DHA. It is worth mentioning that the biocatalysts applications described for lipase in section 1.3.1, such as production of structured triacylglycerol is what figure 4 refers to with 'enzymatic process'.

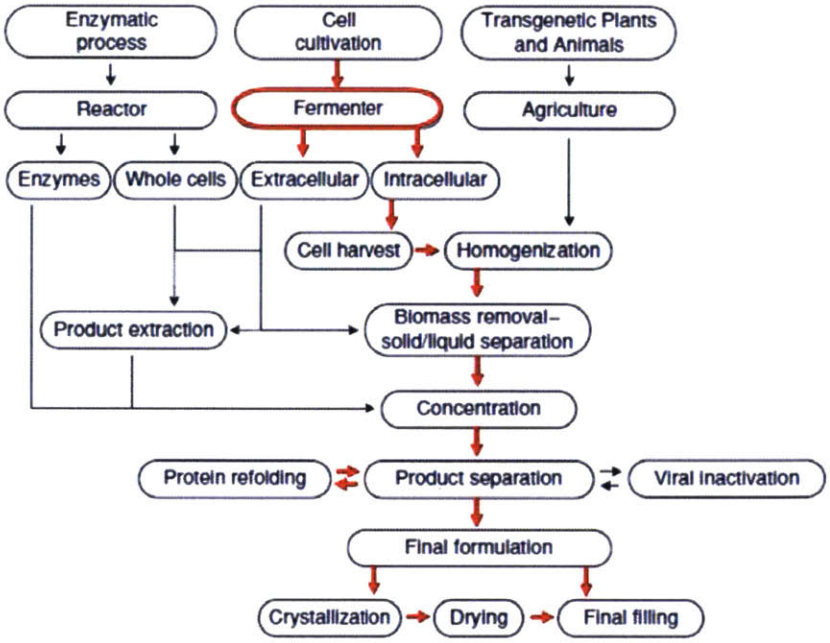


Figure 4 – General applicable process tree for different classes of bioprocess
Red arrows show the process path referred to in this thesis. Red circle highlight the process unit analyzed. Source: Adapted from (Heinzle, Biwer, & Cooney, 2006)

A bioprocess can be divided into 3 sections: Upstream processing, fermentation and downstream processing. "As commonly done in process engineering, we consider unit operations as basic steps in a production process. Typical unit operations in bioprocesses are for example: sterilization, fermentation, enzymatic reaction, extraction, and filtration or crystallization. A unit procedure we define, as a set of operations that take place sequentially in a piece of equipment, e.g. charging of substrate to a fermenter, addition of acid to adjust pH, reaction, transfer of fermentation broth to another vessel" (Heinzle, Biwer, & Cooney, Development of Sustainable Bioprocesses- Modeling and Assessment, 2006). Upstream processing includes all the unit operations that are necessary before the fermentation step. Typical upstream unit operations are: (1) preparation and storage of solutions, (2) sterilization of raw material, (3) inoculum preparation. Fermentation is the main unit operation in bioprocessing, of which the bioreactor is at the core. Fermentation will be reviewed in more detail in the following subsection. Finally there are several possible unit of operations for downstream processing. The selection among all the possible unit operations, specific equipment and corresponding unit procedures is based on the desired properties of the product, the impurities and the microorganism used. The following figure shows typical unit operations for downstream processing and possible techniques for each of them.

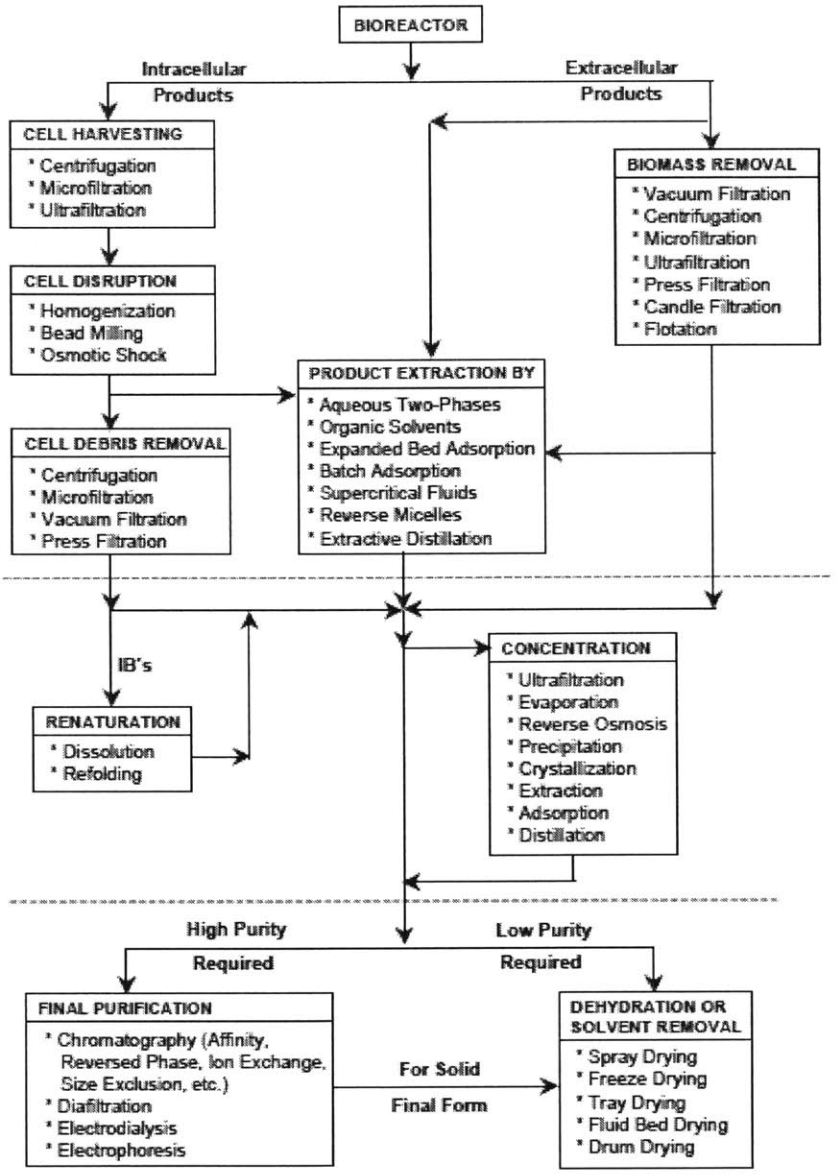


Figure 5 – General block diagram of downstream processing
Source: (Petrides, 2000)

While the author of the present thesis recognizes the importance of analyzing the whole process when developing a Tradespace analysis, (or any cost, competitiveness or sustainability analysis for that matter), due to time constrains, the preset thesis will focus on the fermentation step rather than the full bioprocess of lipase and DHA production. The usefulness of such Tradespace analysis should be then analyzed under such constrain, which will be further discussed in section 5.

Since the scope of this thesis will be the fermentation step of lipase and DHA bioproduction, the following subsection will briefly introduce fermentation technology.

1.4.1. FERMENTATION TECHNOLOGY

Any fermentation is initiated by inoculating a cell culture into the appropriate growth media, a complex mixture that contains the substrate that will be processed by the cell and converted into the targeted product. The fermentation occurs under conditions that favors cell growth and production of the desired compound. The media may contain little or no "free" water, corresponding to a solid-state fermentation, or the substrate can be dissolved (e.g. sugar solution) or suspended in a large amount of water to form a slurry or broth, known as submerged fermentation (Chisti, 2010). Solid-state fermentation is mainly used for filamentous fungi. Thus, the focus of this thesis and the fermentation used in both cases analyzed is submerged fermentation using the most commonly used type of reactor, namely stirred tank bioreactor.

The stirred tank fermenter is one of the most commonly used type of reactors due it its flexibility. It consists of a cylindrical vessel, with a central shaft that typically supports 3-4 impellers. The height-to-diameter ratio of the vessel can vary from 1 to 4. For aerated bioreactors a higher ratio is preferred in order to prolong the contact time between rising bubbles that carry oxigen and the liquid phase. The vessel is typically provided with four equally spaced vertical baffles that extend from the wall into the vessel. The objective of these baffles is to increase the mixing quality. The air, usually supplied by a compressor, enters the vessel at the bottom under pressure. The mixing and bubble dispersion are accomplished by mechanical agitation. This requires a relatively high energy input per unit volume. A jacket and/or internal coils allow heating and cooling. Bellow a schematic diagram of a stirred tank fermenter is shown. (Chisti, 2010; Heinzle, Biwer, & Cooney, 2006)

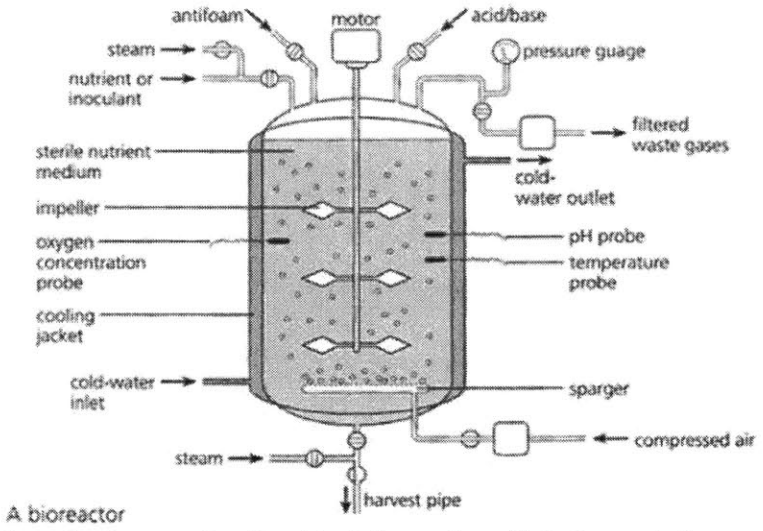


Figure 6 – Schematic diagram of a stirred tank fermenter with instrumentations and controllers
Source: (iGEM2010, n.d.)

Finally, fermentations can be carried out in batch, fed-batch or continuous cultures. Figure 7 shows the difference between these three modalities. The two cases analyzed use the most common process, which is the batch process.

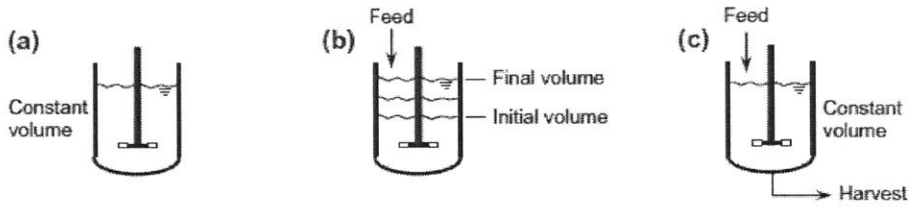


Figure 7 – Fermentation process.
(a) Batch (b) Fed-batch (c) continuous culture. Source: (Chisti, 2010)

1.5. PRODUCT DEVELOPMENT PROCESS (PDP)

The goal of any bioprocess development is the optimization of process parameters to manufacture a product. A product that has a market (or a potential market), satisfies a customer need and has a market size that justifies the investment in the process development. Once the desired product is clearly defined and specified (quality, purity, etc.) the technical aspects of the process development project takes place, as it is the product specifications and cost that establish the goal of the process development (Heinzle, Biwer, & Cooney, 2006). Similar to traditional chemical industry, bioprocess development typically relies on a set of specialist teams, each optimizing with respect to a small set of criteria, e.g. chemical/biochemical synthesis route with respect to product quality and yield, process design with respect to selecting, designing and

connecting suitable unit operations and process control structure, plant design with respect to equipment and process control. The result is a sequential process design with only scarce information transfer between steps (Figure 8.A) (Heinzle & Hungerbuhler, 1997).

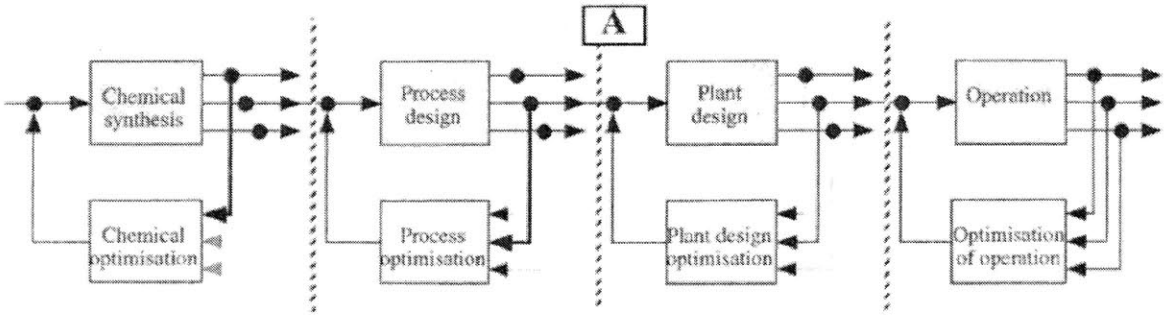


Figure 8 – Information flow in the design process.

Classical sequential process with few interactions between development steps and groups Source: *(Heinzle & Hungerbuhler, 1997)*

The main issue with a sequential process design is that the cost and effort required to correct a sub-optimal decision made at the beginning of the development process increase with time, as the development freedom decreases considerably. However, our knowledge and understanding of a given process is low at the beginning, thus the chances of making a mistake or taking a non-optimal decision due to lack of information is high (Figure 9).

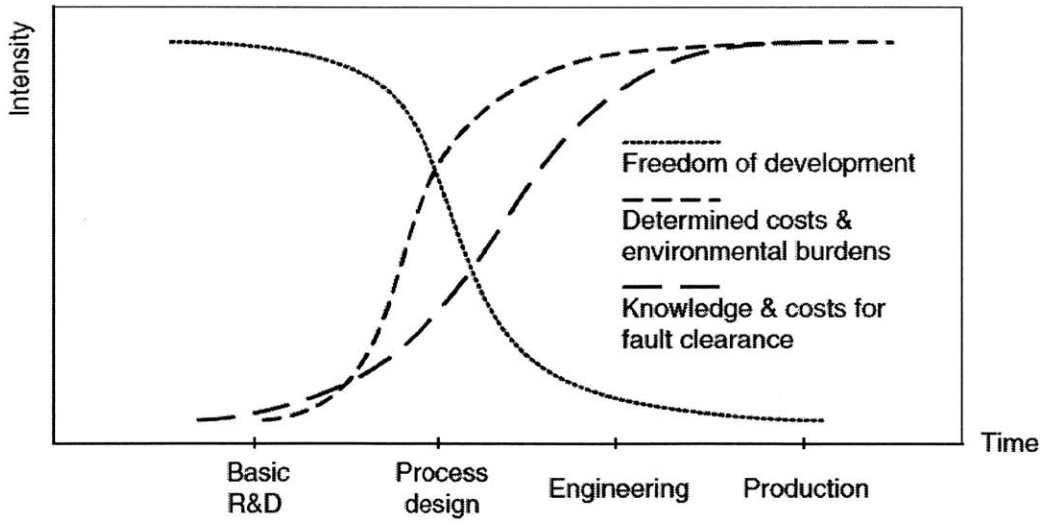


Figure 9 – Process knowledge and freedom of decision in the process design.
Source: (Heinzle, Biwer, & Cooney, 2006)

Since the mid-1980s modeling and simulation software specific to biochemical processes were developed to gain understanding of bioproduction process (More detail in section 2.3). The

use of these software tools help the process design team fill the knowledge and data uncertainty gap and provide a sound evaluation basis. Process simulation software tools also enable the representation and analysis of integrated process. Heinzle et al (1997) defines integrated development (IPD) as a process where:

1) In every development step, information of all other steps is considered, design is done in parallel rather than sequentially (time dimension of IPD);
2) Design alternatives are simultaneously assessed interactively with increasing depth for multiple criteria – economic, safety, and environmental protection (depth dimension of IPD)
3) Impact on local and global environment are simultaneously considered (space dimension of IPD);
4) People with various expertise work in a broad networked multi-disciplinary teamwork (human resource dimension of IPD)

One of the most cited simulators in the literature is SuperPro Designer™ from Intelligen, Inc (Heinzle, Biwer, & Cooney, 2006; Petrides, 2000; Shanklin, Roper, Yegneswaran, & Marten, 2001; Petrides, Carmichael, Siletti, & Koulouris, 2014). This, and other software tools, build the simulation based on a process flow diagram (PFD), which is given as an input, together with process parameters such as scale, operation conditions and performance. Then the simulator performs material and energy balances, cost analysis and economic evaluation (Figure 10). SuperPro Designer™ can also perform tasks such as scheduling and environmental impact assessment, debottlenecking and throughput analysis. However, since operation parameters are given as an input, and alternative process setups can be simulated by "experimenting" on the computer in a trial and error fashion, the scenario analysis of these software is limited. Variations of the process flow diagram, process scale, and operations conditions can be performed only one at a time.

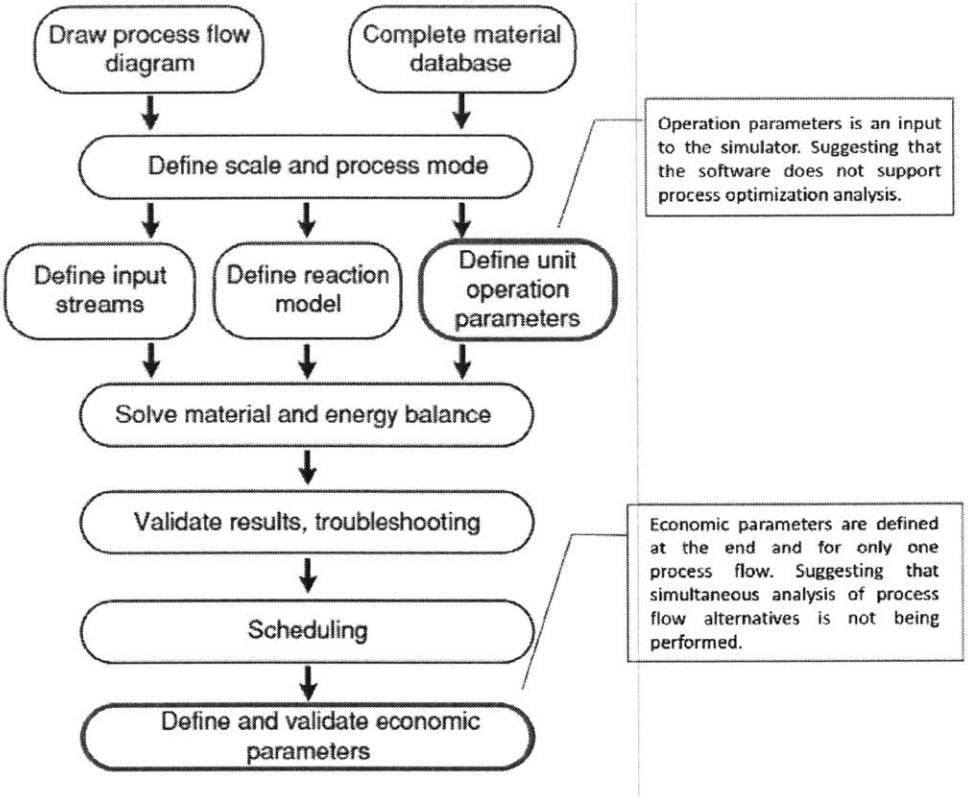


Figure 10 – Steps to build a model in process simulation software.
Source: (Heinzle, Biwer, & Cooney, 2006)

As Heinzle et al. explained in chapter 3 of his book (Development of Sustainable Bioprocesses- Modeling and Assessment, 2006), "especially in early process development, there might be a need to compare alternative process flow sheet topologies. An extraction step might replace a distillation column or the downstream steps might vary. For such changes the economic and environmental impact can be derived in a scenario analysis. Furthermore, variation in size and number of pieces of key equipment, namely the fermenter, can be studied with scenarios". However, current software tools perform this analysis by creating new files for each new scenario, and comparing each scenario to the base model one by one, as shown in the example bellow.

Table 5 –Scenario analyses of the cellulose production model.
Table 5 shows the result of two scenario analysis for the cellulose model. In the base model, the inoculum volume is 5% of the fermenter volume. This defines the necessary volume of the seed reactors. If the inoculum volume is increased, the starting cell concentration is higher, and thus the time to reach the maximum biomass concentration and product formation might be shorter. In this scenario we assume the fermentation time to be 10h shorter when the inoculum volume is increased by 10%. This enables a higher annual production. However, it requires an increase in the size of the seed reactors, which causes higher investment cost. This additional cost outweighs the higher annual production and causes higher unit production cost.
The second scenario describes the situation when an additional ion-exhange adsorption step is necessary to remove some interfering by-products. This additional step not only raises the investment cost but also reduces the annual production (product loss).
Source: (Heinzle, Biwer, & Cooney, 2006)

Scenario	Annual production (metric tons)	Capital investment ($ million)	Unit production cost ($/kg cellulase)
Base case	456	20.7	15.4
10% Inoculum	475	23.4	16.4
Additional chromatography	385	22.1	20.4

By setting a baseline "favorite", or previously developed concepts, and performing scenario analysis 'one at a time' can lead to premature reduction of topologies or architectures. Premature focusing can introduce artificial constrains on the design process and reduce potential value created and delivered (Ross & Hasting, 2005). Other industries, such as aerospace, have addressed the possibility of comparing several topologies/architectures using methodologies such as Tradespace analysis.

1.6. TRADESPACE BASICS

Ross and Hasting (The Tradepace Exploration Paradigm, 2005) define Tradespace as "the space spanned by the completely enumerated design variables, which means given a set of design variables, the Tradespace is the space of possible design options". In general, a Tradespace is a representation of a set of architectures in a space defined by two or more metrics. It differs from the method used in example shown in table 5, usually referred to as 'Point-Based Design' (Bernstain, 1998), in that Tradespace allows the designer (and other decision-makers) to explore the design space taking into consideration a set of strongly interdependent variables, and optimize for more than one metric/objective. The most common objectives are maximing performance and minimizing cost. In figure 11 bellow, a schematic Tradespace plot is shown. For a given a cost and performance threshold, the optimal area (in blue) and along the curve ("Pareto frontier") are the potential solution that balances cost, performance and schedule (known in project management as the 'Triple Constrain' or 'Iron Triangle').

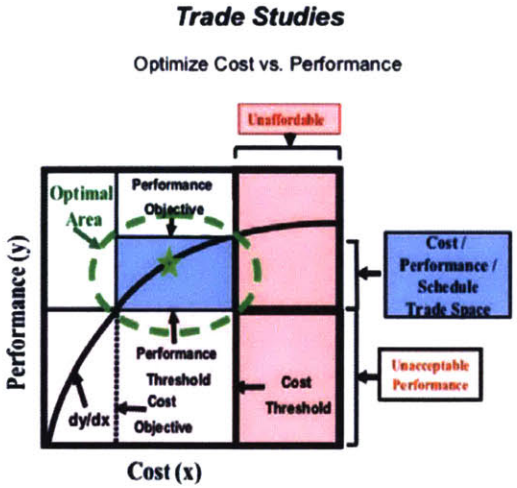


Figure 11 – Schematic representation of a typical Tradespace plot
Source: (Mackertich & Kraus, 2008)

1.7. SUMMARY AND THESIS DESIGN

In section 1 I have introduced the context of this thesis by establishing an overview of industrial biotechnology, its importance and applications in food and feed, as well as the general economic and application within food and feed for the two case studies chosen for further analysis. Also, a description of bioprocesses, especially fermentation technology and biocompounds PDP, was presented. Finally a general introduction to Tradespace and how it addresses the issue of premature focusing by allowing an extensive architecture exploration, not possible with current process simulation software, was introduced. Thus, the objective of this thesis is to explore the use of Tradespace in the context of bioprocesses by developing a Tradespace focused in the fermentation step using lipase and DHA production as case studies. Furthermore, how this Tradespace fits into the larger context of a broader analysis involving the full bioprocess and a conceptual design of such Tradespace is presented.

In order to build a Tradespace for the production of the selected biocompounds, the present thesis combines methods commonly used in bioprocess optimization methods, such as Design of Experiments (DoE) - particularly the use of Plackett–Burman designs and Response Surface methodology- with an existing Tradespace exploration method known as Multi-Attribute Tradespace Exploration (MATE). These concepts will be further reviewed in the next section, 'literature review'.

2. LITERATURE REVIEW

2.1. MATE- MULTI ATTRIBUTE TRADESPACE EXPLORATION

2.1.1. MATE HISTORY

The Multi-Attribute Tradespace Exploration (MATE) method was develop at the Massachusetts Institute of Technology by Adam Ross, Nathan Diller, Dr. Dan Hasting, Dr. Joyce Warmkessel, Dr. Hugh McManus, and others (Spaulding, Tools for Evolutionary Acquisition: A Stuydy of Multi-Attribute Tradespace Exploration (MATE) Applied to the Space Based Rada (SBR), 2003). MATE's development began with system analysis work done in the MIT Space System Lab, which was eventually embodied in a process called Generalized Information Network Analysis (GINA). GINA's goal was to model satellites as information networks, and it used metrics (appropriate for information systems) to construct a Tradespace of possible designs. Over the years GINA was applied in a series of aerospace related projects, and in each iteration new methods were included, evolving eventually into the MATE method (Spaulding, 2003). MATE combines two techniques used in technical design and decision making: Multi-Attribute Utility Theory (Keeney & Raiffa, 1993) and Tradespace Exploration (Ross & Hasting, 2002).

The present thesis incorporates DoE within the MATE method in order to develop a Tradespace process for bioproduct production. The reason to incorporate DoE will be discussed in sections 4 and 5. An overview of the MATE process at a level of detail appropriate for general understanding will be introduced in the next subsection. A parallel of how some of the concepts/terms used in the MATE process can be translated into the context of a bio-compound production will be discussed in the Tradespace model presented in section 4. Once the overall MATE method is discuss, in the subsequent subsections, Multi-Attribute Utility Theory (MAUT), Tradespace Exploration, and DoE will be further introduced. Finally, an overview of existing process simulators will be provided as a background for the discussion presented in section 5.

2.1.2. MATE PROCESS OVERVIEW

The MATE process consist of the following steps:

1. Identify stakeholders
2. Define a mission objective/concept
3. Create a list of attributes
4. Determine design variables and map them to the attributes

5. Create a model that gives rise to utility curves
6. Evaluate architecture

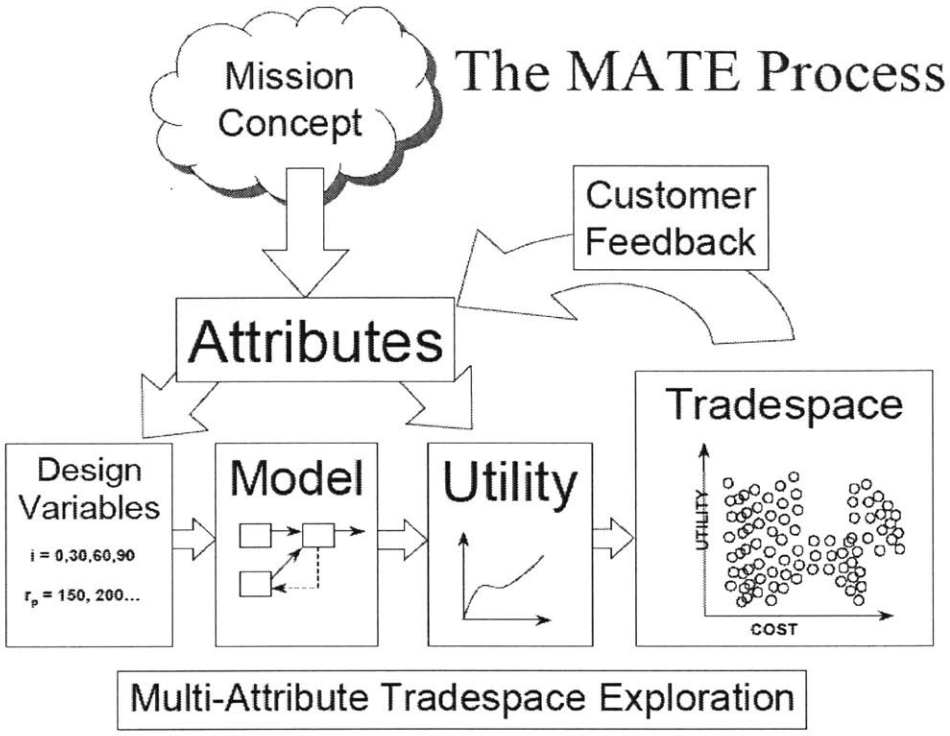


Figure 12 – The MATE process.
Source: (Spaulding, 2003)

In order to better understand each of these steps, it is helpful to define the concepts/terms used. Some of these terms are part of the general systems engineering lexicon, however others are unique to MATE or have a more specific definition when applied to MATE. The following list is adapted from (Ross & Hasting, 2005), (Spaulding, 2003) and (Crawley, Cameron, & Selva, 2016):

Mission Concept/Objective: the desired end state or outcome of the system.

System: a set of entities and their relationship whose functionality is greater than the sum of the individual entities.

Attribute: a metric perceived by a decision-maker that measures or determines how well the defined objective is met.

Utility: a dimensionless parameter that reflects the 'perceived value under uncertainty' of an attribute. Typically ranging from zero (minimal acceptance) to one (most desirable). Utility is a useful quantitative proxy for representing benefit.

Multi-Attribute Utility: a dimensionless parameter, ranging from zero to one, that reflects the value of an aggregation of single utility values.

Design Variable: a designer-controlled parameter.

Architecture: a potential system – a unique combination of design variables

Tradespace: the set of all architectures under consideration

With these definitions in mind, an interpretation to how each of these concepts/terms can be applied in the context of a bioprocess is defined in section 4.

2.1.3. MULTI-ATTRIBUTE UTILITY THEORY

Multi-Attribute Decision Making (MCDM) is the most well-known branch of decision making. It is a branch of a general class of Operations Research models. MCDM combines theories from disciples such as philosophy, mathematics, and psychology that try to explain and formulate the logic behind decision making (Nikou & Klotz, 2014). Decision theory is widely applied in economic, mathematics and social sciences (Raiffa, 2002). MCDM is divided into: (a) Multi-Objective Decision Making and (b) Multi-Attribute Decision Making. Multi-Objective Decision Making studies decision problems in which the decision space is continuous. On the other hand, Multi-Attribute Decision Making concentrates on problems with discrete decision spaces. In these problems the set of decision alternatives has been predetermined (Triantaphyllou, Shu, Nieto Sanchez, & Ray, 1998). Several methods exist for Multi-Attribute Decision Making, one of which is MUAT (Multi-Attribute Utility theory).

MUAT was developed by von Neumann and Morgenstern (1944/1947/1953), and it was referred to as the "Expected Utility Hypothesis". MAUT makes possible the calculation of overall utility (i.e. customer satisfaction or preference) of multiple attributes (i.e. product characteristic or features) based on single utility functions. Single-Attribute Utility (SUF) is a dimensionless metric representing the satisfaction derived from having a certain level of a single attribute X. Single-Attribute Utility function can be used to express the relative desirability of having a specific value of an attribute. In equation:

$$SUF_i(x_{ij}) = u_i(x_{ij})$$

Where SUF is single utility function for attribute i, x_{ij} is the attribute rating (the raw score) for alternative j of the attribute i, $u_i(x_{ij})$ is the utility function that transforms the attribute rating into a utility value between 0 and 1. Linear utility functions are used to model each attribute. More

sophisticated approaches are available in the form of non-linear utility functions, but should be used only if there is explicit need to include non-linear behavior (Ogle, Dee, & Cox, 2015).

To compare designs that have more than one attribute of interest, SUFs need to be combined into a multi-attribute utility function U. One of the simplest forms of multi-attribute functions is the weighted sum:

$$U(j) = \sum_i w_i u_i(x_{ij})$$

$$\sum w_i = 1$$

Where $U(j)$ is the Multi-Attribute Utility Function (MUF) corresponding to the j^{th} alternative. Multi-Attribute Utility is the joint utility level derived from multiple attributes.

As explained above, the MATE method combines MAUT with Tradespace analysis. While the MUAT component of MATE allows for a systematic assessment of different multi-attribute designs, Tradespace provides a visual tool to compare these designs.

Ross (2003) justifies the choice of MAUT in MATE to capture user preference in the following way:

> "It [MAUT] provides for a systematic technique for assessing customer "value", in the form of a preference for attribute. Additionally, it captures risk preferences for the customer. It also has a mathematical representation that better captures the complex trade-offs and interactions amongst the various attributes. In particular, the strength of Multi-Attribute Utility Analysis lies in its ability to capture a decision maker's preference for simultaneous objectives"

2.1.4. THE PARETO FRONTIER AND TRADESPACE STRUCTURE

The Pareto frontier, is simply a set of architectures that form the 'edge' of the Tradespace. Because we have two or more metrics represented in a Tradespace, it is unlikely that any single architecture is uniquely "the best". Rather, the Pareto frontier or Pareto front, showcases the architectures that are "good" and represent a good tradeoff between the metrics (Crawley, Cameron, & Selva, 2016).

Pareto analysis in an important part of Tradespace analysis, but it is by no means sufficient. A lot can be learned by analyzing the structure of the Tradespace as a whole. Often times a Tradespace looks like a "cloud" of points. However, most Tradespaces have some structure. They

have features such as "holes", "subgroups" and "fronts". These are due to factors such as discrete metrics, different dynamic ranges of metrics and physical laws limiting certain metrics. (Crawley, Cameron, & Selva, 2016)

A common structure in Tradespace are "clusters"- that is, the accumulation of architectures in relatively small regions in the objective space, leaving large open areas, Fig 13. Clusters suggest the presence of families of architectures that achieve similar performance in one or more metrics. It is useful to view clusters with similar architectural variables, which can simply be done by highlighting the points that share the same decision choice.

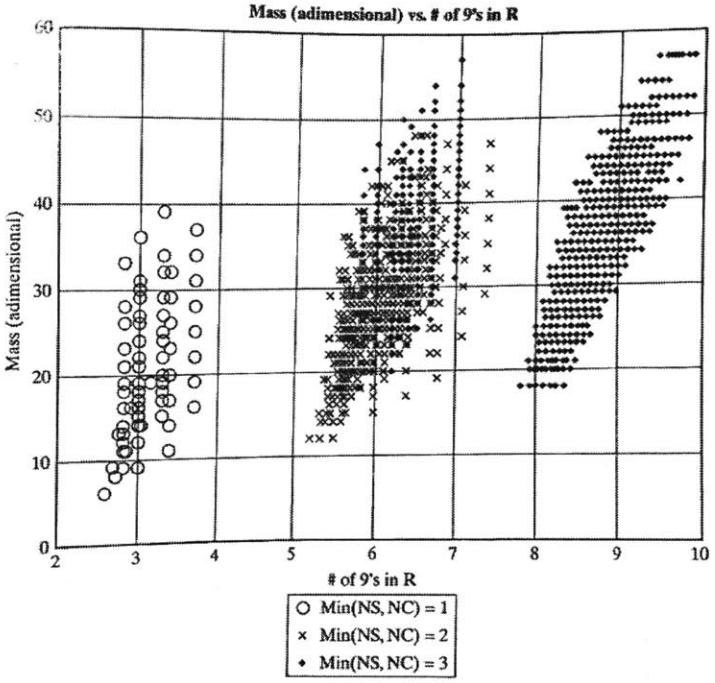


Figure 13 – Tradespace showing 3 clusters
The example shows a mass-reliability Tradespace for GNC system (Guidance, Navigation, Control system, present in vehicles, such as cars, aircraft, robots, and spacecraft). Architectures are highlighted with different markers depending on their numberof sensors (NS) and the number of computers (NC). Where min(NS, NC) =1 (circles) means that the architectures there have either one sensor or one computer. Min(NS,NC)= 2 (crosses) are architectures with at least a sensor and a computer, two sensors or two computers. Similarly min (NS, NC) = 3 is the min number of computer, sensors and/or their combinations. The number of nines (# of 9's) in the realiability is shown instead of the reliability value (for example 3 nines is equivalent to R=0.999 and R=0.993 is equivalent to 2.15 nines). Source: (Crawley, Cameron, & Selva, 2016)

Another common feature in Tradespace is stratification. Strata are groups of point for which one of the metrics is constant while the other varies. In a two-dimensional Tradespace, strata appear when points line up in a number of vertical or horizontal lines, which gaps in between them.

Stratification appears when combinations of architectural options produce only few distinct values of the metric.

2.2. DESIGN OF EXPERIMENTS (DOE)

Statistical experimental planning, factorial design and Design of Experiments (DoE), are more or less synonymous concepts for investigating the mathematical relationship between input and output variables of a system. Even though the fundamentals of the methodology have been known since the early 1900s, it was not until the late 1990s that is was widely applied in biotechnology (Mandenius & Brundin, 2008). When used to optimize processes, DoE is a systematic way of changing process inputs (e.g. temperature, pH, medium components, etc.) and analyzing the resulting process outputs (e.g. yield, productivity, etc.) in order to quantify the cause and effect relationship between them, as well as the random variability of the process while using a minimum number of runs. The conventional approach to optimization investigates One Factor at A Time (OFAT) while keeping the others constant. Unlike OFAT, DoE detects the interaction between input parameters, requires less number of experiments and facilitates the prediction of the response to values not yet tested in the experiment. By performing factorial design, a reliable result can be achieved with relatively fewer experiments, after which the most favorable direction to move forward in order to find a true optimum can be evaluated. For example, in figure 14, the diagram to the left explains the OFAT approach to optimize/investigate in a three-dimensional parameter space, where the parameters/factors X, Y and Z can be changed, one at a time, giving different outputs. As represented in the figure, the OFAT approach does not cover the complete three dimensional space. On the other hand, on the right, a full-factorial three-parameter DoE is represented. By analyzing all the "corner points" and "mid point" of the three dimensional space, the complete design space is studied.

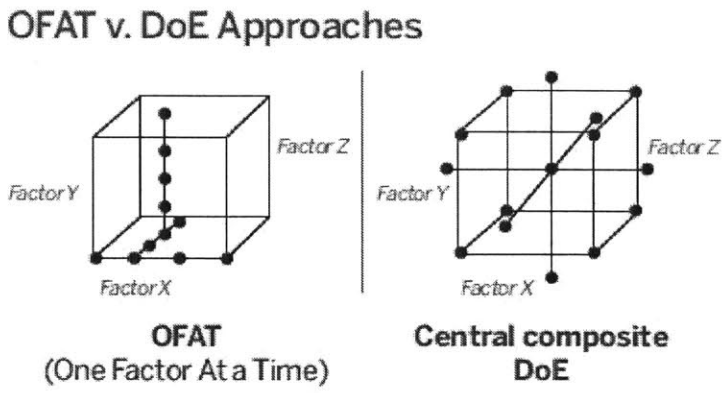


Figure 14 – OFAT v. DoE approaches
Source: (Owen, et al., 2001)

Furthermore, through DoE systematic and unsystematic variability is also studied. Thus, DoE requires fewer resources for the amount of information obtained, especially as the number of factors increases.

In DoE, the dependent variable (process output) is called response and the independent variable (process inputs) are called factors. The simplest factorial experiment has two factors, with two levels each. The annotation for factorial experiments is: X^y; where X= #levels and Y= #factors. Thus, the simplest factorial experiment is a 2^2, producing 4 factorial points. Figure 14, represents a factorial design of 2^3: 3 factors, 2 levels each factor and thus 8 tested conditions. When all the possible combinations across of levels across the studied factors are tested, it is known as a Full Factorial Design. However, as the number of factors studied increases, carrying out the experimental design might become logistically unfeasible. In this case, a Fractional Factorial Design can be performed. For a two level DoE, the reduced set of experiments can be described mathematically as 2^{n-k}, where n is the number of factors and k the number of steps to be reduced. If for example five variables are involved in the experiment, we will end up with 16 (2^{5-1}) or 8 (2^{5-2}) experiments, depending of the number of steps reduces. In practice, the number of steps reduced and the set of combinations chosen to test (known as principal fraction), are described in statistical reference books and "standard" factorial design exists and are chosen depending on the objective of the study (Box, Hunter, & Hunter, 2005).

Generally speaking, there are usually 3 types of objectives for a DoE: (1) screening, (2) optimizing and (3) robustness testing. Two experimental designs were used by the authors of the papers used as resource for the case studies in this thesis. The first for screening, namely Plakett-

Burman experimental design and the second Response Surface Methodology (RSM) for optimization. Both designs are further discussed in the following subjections.

2.2.1. PLACKETT-BURMAN EXPERIMENTAL DESIGNS

Plackett-Burman (PB) designs (Plackett & Burman, 1946), is a fraction of a two-level factorial design and allows the investigation of *n-1* factors in at least *n* experiments. PB designs are well suited to establish whether the outcome of an analytical procedure is affected by changes in each relevant factor. They have become known for their ability to investigate a large number of factors in a relatively low number of experimental runs. This becomes possible, because the interactions between factors are neglected in PB designs, thus they are very efficient in screening factors when only main effects are interested.

The number of runs *n* in a PB design are multiples of four. Plackett and Burman only included designs with $n \leq 100$, and omitted the design where $n = 92$. In each case the maximum number of factors that can be studied is n-1, so an 8-experiment PB design can study no more than 7 factors, a 12-experiment design will handle up to 11 factors, and so on. Note that the number of runs must be a multiple of 4, therefore, if 4 factors are studied, that is n-1=4, n=3. But 3 is not a multiple of 4, thus the number of runs needed to study 4 factors in a PB design is 8 experiments and 7 factors. There are 4 factors of interest and 3 "dummy" factors. Dummy factors are those that have no physical significance, but will nevertheless inform about random measurement errors. Usually, three dummy variables will provide an adequate estimate of errors (Stowe & Mayer, 1966).

As mentioned above, PB designs use two levels for each factor. The higher level is denoted as '+' and the lower '-'. A further feature of the PB method is that the + and – signs for the individual run are assigned in a cyclical manner. For example, for a seven factor experiment labelled A-G, the first experiment might be (Miller, 2013):

A	B	C	D	E	F	G
+	-	-	+	-	+	+

The levels for the second experiment, again with four + and three –, are obtained by shifting the sigs of the first experiment one place right and moving the last sign to the begging of the line, giving:

A	B	C	D	E	F	G
+	+	-	-	+	-	+

This cyclical process is repeated for the first seven experiments. For the eighth experiment all factors are set at the low (-) level, giving an overall design in which there are 28+ signs and 28- signs, each factor having been studied four times at the high level and four times at the low level. In practice, the sequence of + and – signs are provided by generating vectors and are widely available in the literature and software packages. Some commercially available software packages frequently used for experimental designs are Modde™ (Umetrics AB, Umeå, Sweden; www.umetrics .com), MiniTab™ (Minitab Inc., State Collage, PA) and Design-Expert™ (www.statease.com), all of which are convenient for applying DoE (Mandenius & Brundin, 2008).

The effect of each factor is then determined from the expression:

$$E_{(xi)} = \frac{\sum M_{i+} - \sum M_{i-}}{N/2}$$

Where $E_{(xi)}$ is the main effect of the tested variable *Xi*. (M_{i+}) is the response when factor *Xi* is at its high level and (M_{i-}) response when a same factor *Xi* is at its low level. *N* is the total number of experiments.

The significance of each factor is then calculated through analysis of variance (ANOVA) related calculations. As an example, if we consider that the sum of the squares (SS) is given by (Miller, 2013):

$$\text{SS} = N \times (\text{estimated effect})^2/4$$

In the previous example, for each factor, the sum of squares only has one degree of freedom, so their mean value (i.e. variance) is the same as their SS. For the dummy variables however, the mean of sum squares is calculated depending on the number of dummy variables used; in this example 3. Then each individual factor can be compared with the estimated random error (mean sum square of dummy variables) using a one-tailed *F-test* at p=0.05 (95% confidence level). If we consider that each factor has one degree of freedom and the dummy variables have three degrees of freedom (three dummy variables were used), $F_{1,3}$ at p=0.05 is 10.13 (Appendix B). Thus, for each factor

$$F = \frac{mean\ square\ of\ factor}{mean\ square\ of\ dummy\ variable}$$

F> 10.13 then factor under evaluation is significant

F <10.13 then factor under evaluation is NOT significant

Once again, such calculations are in practice performed using suitable software.

2.2.2. RESPONSE SURFACE METHODOLOGY (RSM)

The result of a bioprocess screening experiment, using a factorial DoE such as PB designs, is the identification of a subset of most influential factors. These factors can be used in a new experimental design with the purpose of determining optimal factor values. The experimental results of such DoE lead to the deduction of a function that can explain the response:

$$y = f(x_1, x_2) + \varepsilon$$

Where ε represents the noise or error observed in the response y. The surface represented by $f(x_1, x_2)$, is called a response surface.

A response surface can be represented graphically, either in the three dimensional space or as a contour plot that helps visualize its shape.

An experimental design commonly used in RSM is Central Composite Design (CCD) (Montgomery, 1997). The CCD is a very effective design for fitting second order response surface, where the behavior of the system can be explained by the following quadratic equation:

$$Y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} x_i^2 + \sum \beta_{ij} x_i x_j$$

Where Y is the predicted response; β_0 the offset term, β_i the linear effect, β_{ii} the squared effect and β_{ij} the interaction effect.

Central composite designs are factorial or fractional factorial designs with center points, augmented with a group of axial points (also called star points, represented in blue in figure 15) that allows the estimation of curvature. Figure 15 illustrates a three variable case of CCD.

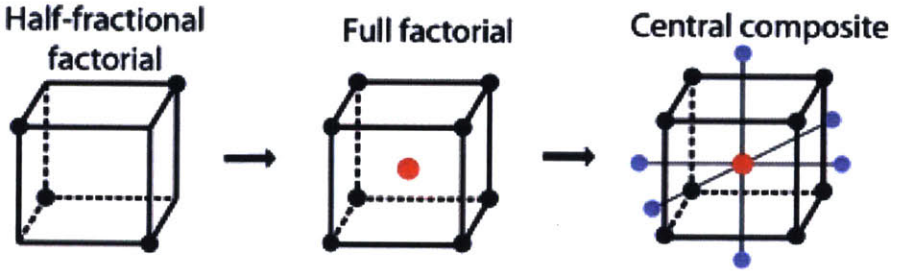


Figure 15 – Central Composite DoE
Source: (Lebed, Potvin, Lariviere, & Dai, 2014)

There are three types of CCD:

1) *Central Composite Circumscribed* (CCC): in a three factor CCC design, the high (+1) and low (-1) levels are represented at the corner of the cube. Star points are displaced outside the space at the same distance from the center point as the distance from the center to the corners. (Fig 16.A)
2) *Central Composite Face-centered* (CCF): In a three factor CCF design, star point are located between the high and low level. Thus, between +1 and -1. (Fig 16.B)
3) *Central Composite Inscribed* (CCI): in a three factor CCI design, the star points take the values + 1 and -1, while the high and low levels lie inside the in the interior at the corners of the cube. (Fig 16.C)

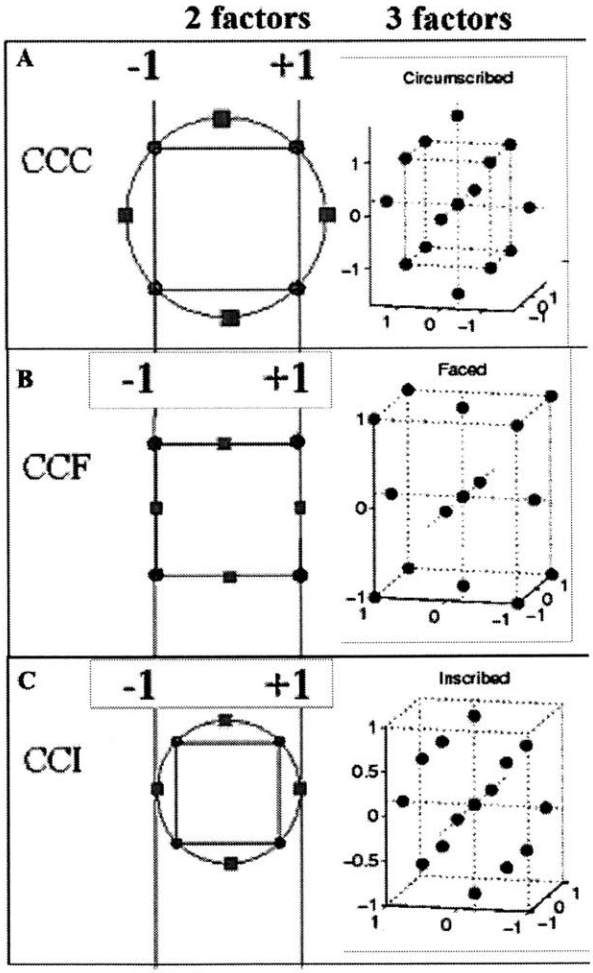


Figure 16 – Comparison of Three types of Central Composite DoE
Source: Modified from (NIST/SEMATECH, 2013)

For the two case studies analyzed in this thesis different central composite designs were used. The lipase case study uses a CCF design, whereas the DHA case study uses a CCC design. In theory, the CCC design is somewhat better than the CF design since CCC covers a larger volume. Also, in CCC a total of 5 levels per factor are evaluated. This is because the star point are displaced outside the space (usually coded as +2 and -2) and therefore represents two additional levels. These two new levels plus the original 2 levels (+1 and -1) and the center point result in a total of 5 levels. In contrast CCF only allows for 3 levels to be evaluated per factor (-1, 0, +1). Thus, CCC can therefore better capture strong curvature and even cubic responses.

2.3. BIOPROCESS SIMULATION SOFTWARE

Process Simulators are software tools that enable the representation and analysis of integrated processes. Process simulation tools were first implemented in chemical and petrochemical industries in the early 1960s. Established simulators for the petrochemical industries include: Aspen Plus (from Aspen Technology, Inc. https://www.aspentech.com/)), ChemCAD (from Chemstations, Inc. http://www.chemstations.com/), HYSYS (developed by Hyprotech, Ltd, acquired by Aspen Technology is 2002), and PRO/II (from Simulation Sciences, Inc., now Schneider-electric http://software.schneider-electric.com/simsci/) (Petrides, Bioprocess Design, 2000). However, these simulators were designed to model steady-state (continuous) processes, therefore they did not account for the sequential nature of batch processes, where a sequence of time-dependent tasks may take place in a given unit operation. The first batch process simulator was named BATCHES. This software was commercialized in the mid-1980s by Batch Process Technologies, a Purdue University spin-off headquartered in West Lafayette, IN (http://www.bptechs.com/). All its operation models are dynamic and simulation always involved integration of differential equations over a period of time. In the mid-1990s, Aspen Technology (Burlington, MA, USA, https://www.aspentech.com/) introduced Batch Plus (later renamed Aspen Batch Process Developer), a recipe-driven simulator that targeted batch pharmaceutical processes. Around the same time, Intelligen (Scotch Plains, NJ, USA, http://www.intelligen.com/) introduced SuperPro Designer. SuperPro, originally BioPro Designer, was initially developed at the Biotechnology Process Engineering Center (BPEC) at MIT. It was licensed to Intelligen, Inc., who completed the development and commercialized it. SuperPro, is an extension of BioPro, and was created to extend its scope to support modeling of fine chemicals, pharmaceuticals, food

processing, consumer products and other types of batch/semi-continuous processes (Petrides, Carmichael, Siletti, & Koulouris, 2014; Petrides, Bioprocess Design, 2000)

While it is beyond the scope and intention of this thesis to explain how process simulators such as SuperPro work, a quick overview is provided in Appendix A. However, it is important to highlight once again, that the purpose of a Tradespace analysis is very different to what these software can provide. A Tradespace objective is to perform what in bioprocess simulations software is known as scenario analysis. However, unlike scenario analysis, where each scenario is created by changing one parameter at a time from the baseline design, a Tradespace analysis allows the analysis of several scenarios at a time. In this way a potential premature focusing, where introduction of artificial constrains on the design process could be avoided.

3. RESEARCH METHOD AND APPROACH

3.1. RESEARCH SCOPE

As explained in section 1.4, while the author of this thesis recognizes and understands the importance of analyzing the whole process, due to time constrains, the Tradespace developed in this thesis only focuses in the fermentation step. However, how this Tradespace might fit into a broader analysis taking into account all 3 steps (upstream process, fermentation and downstream process) will be conceptually discussed.

Furthermore, it is also important to mention that the fermentation Tradespace process suggested in this thesis was analyzed only under the context of submerged fermentation in batch mode using a stirred tank bioreactor. While it theory, the proposed process should be applicable to other types of fermentation (solid-state), it is beyond the scope of this thesis to analyze the applicability of the proposed process under these other settings.

3.2. DATA GATHERING AND ASSUMPTIONS

Due to time constrains, laboratory experiments were not performed throughout this thesis. Instead two published papers were used as basis for the Tradespace design developed. Also, since bioreactor physical details are usually not presented as part of scientific journal articles, the physical characteristics of a 14L New Brunswick bioreactor were used to estimate the electric power requirements. In the following subsections a summary of both papers used as case studies are presented and the physical detail of 14L New Brunswick bioreactor is described.

On the other hand, utility and raw material cost information was collected from various resources.

3.2.1. Lipase Case Study

Lipase fermentation statistical optimization data used in this thesis is based on the study performed by Rathi et al (Statistical medium optimization and production of hyperthermostable lipase from Bukholderia cepacia in bioreactor, 2002). Since an experimental design to determine most relevant attributes/factors was not performed in this paper, the relevance of the factors used in response surface methodology was corroborated through literature review from optimization studies from other *Burkolderia* strains grown in similar conditions. The response surface experimental design and experimental result were used as input for the cost model and the regression equation reported was re-assessed and used as attribute model.

3.2.2. DHA Case Study

Unlike lipase fermentation case study, DHA production parameters relevance were studied through Plackett-Burman experimental design and reported by the author. The source of the data used for the DHA statistical optimization is the study performed by Song et al (Optimization of fermentation parameters for the biomass and DHA production of Schizotrium limacicum OUC88 usign resposnse surface methodology, 2007). The response surface experimental design and experimental result were used as input for the cost model and the regression equation reported was re-assessed and used as attribute model.

3.2.3. 14L New Brunswick bioreactor Characteristics

In order to estimate the electric power requirement in the DHA case study, it was assumed that a 14L New Brunswick bioreactor was used. Also, some constants such as media density and viscosity were also assumed. The physical parameters and constants assumed for this exercise are shown in the figure and table below.

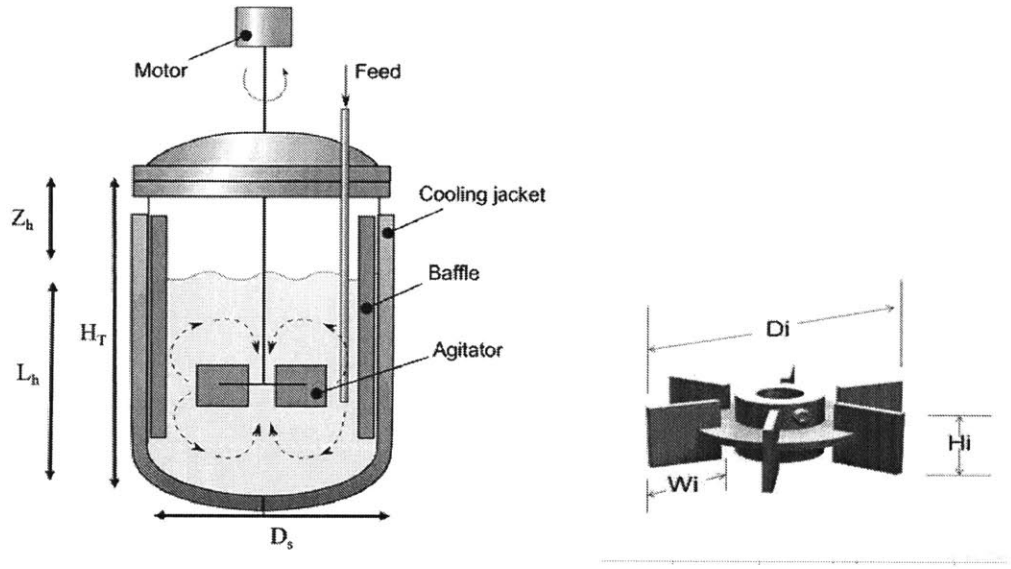


Figure 17 – Schematic diagram of a stirred tank bioreactor and impeller

Table 6 – 14L New Brunswick bioreactor physical parameters and constants used for electric power estimation

Physical Parameter	Symbol	Unit	Value Case Study 1 (DHA)	Value Case Study 2 (Lipase)
Tank				
Tank Diameter	D_s	m	0.211	0.211
Tank total volume	V_t	L	14	14
Tank total height	H_T	m	0.424	0.424
Height of liquid in tank	L_h	m	0.225	0.31
Height of head space	Z_h	m	0.199	0.114
Working volume	V_Z	L	7	10
Volume of dish	V_d	L	2.073	2.073
Volume of head space	V_h	L	6.958	3.986
Cross sectional area of tank	A_t	m^2	0.035	0.35
# baffles	# baffles	#	4	4
Baffle width	b_t	m	0.02	0.02
Impellers				
# of impellers	# of impellers	#	3	3
Diameter of impeller	D_i	m	0.076	0.076
Width of impeller	W_i	m	0.018	0.018
Height of impeller blade	H_i	m	0.015	0.015
Constants				
Media density	ρ	kg/L	1220	1220
Media viscosity	μ	Kg/(s*m)	0.001	0.001
Power number	P_0	-	5.5	5.5

3.3. POWER DENSITY – EQUATIONS

The power density (P/V) was used to estimate electric power requirements in the cost model. Electric power is expressed in terms of power per unit volume in Watts per cubic meter (W/m^3). The power consumption for an un-aerated reaction mixture in a stirred tank reactor is define by the equation (Holland & Chapman, 1966)

$$\frac{P}{V} = \frac{P_o \rho N^3 D_s^5}{\pi D_v^2 L_h / 4}$$

Where P_0 is the power number of the impeller, N, the impeller rotational speed (rps), D_s, the impeller diameter (m), D_s, the vessel diameter (m) and L_h the height of the liquid in the reactor (m).

3.4. COST MODEL AND INITIAL TRADESPACE

Cost models were built on Microsoft excel as well as the regression models re-assessment and Tradespace plots. For the regression model re-assessment Solver add-in (included with excel 2010, 2013 and 2016) was used. Lipase and DHA experimental results were used as input to perform a non-lineal model fitting in excel. The steps to perform a non-lineal model fitting in excel are as follow (Figure 18):

Step 1- Create a table with the central composite rotatable design including the experimental results

Step 2- Assign random values to the coefficients in the quadratic equation.

Step 3- Add a column for predicted results and introduce the corresponding quadratic equation

Step 4 – Calculate the squared error and sum of squared errors

Step 5- Use solver to determine the minimum sum of squared errors (objective function) by changing the coefficients

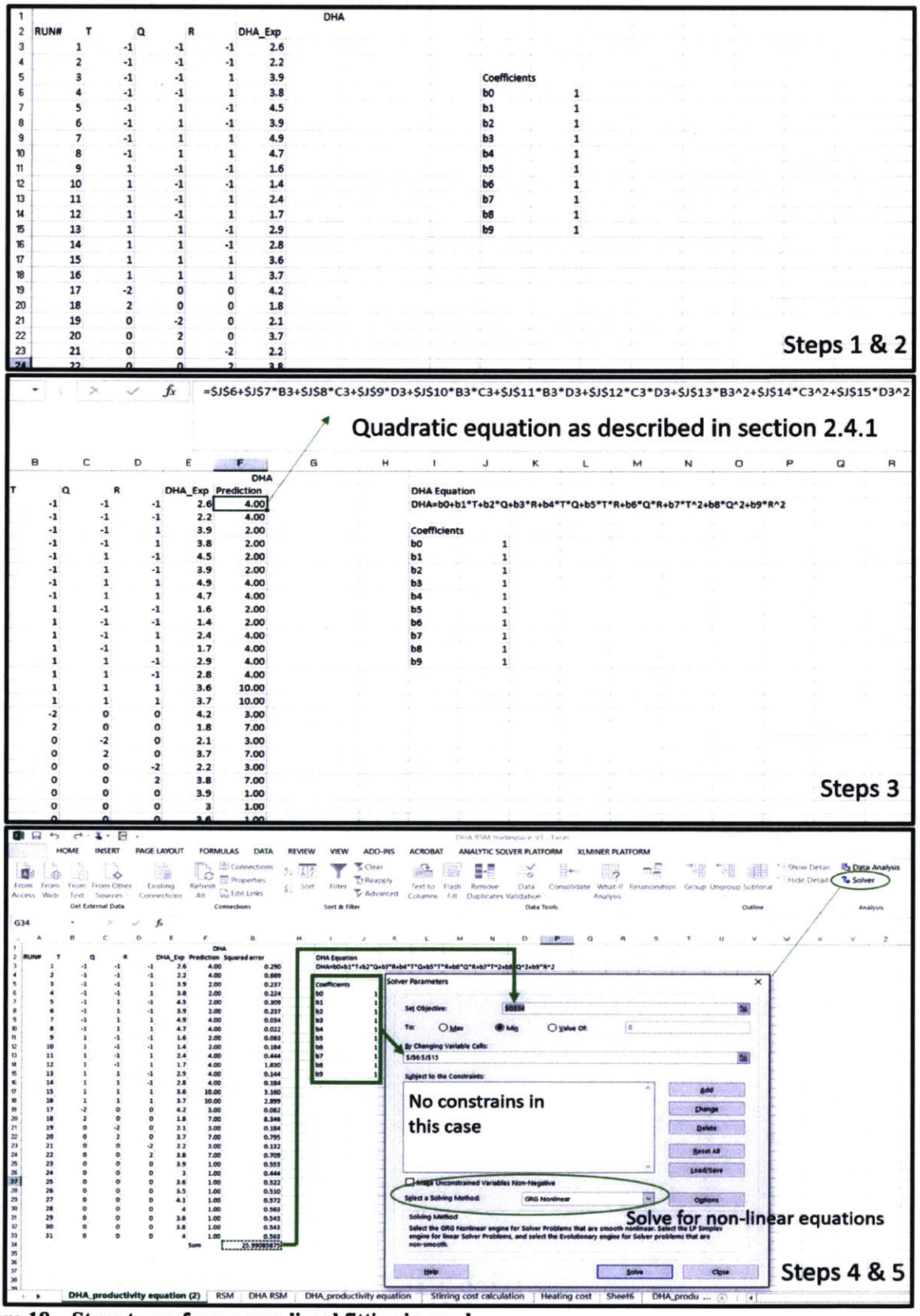


Figure 18 – Steps to perform a non-lineal fitting in excel

4. TRADESPACE MODEL FOR A BIOPROCESS FERMENTATION

4.1. INTRODUCTION TO TRADESPACE CASE STUDY

4.2. MATE APPLIED TO BIOPROCESS

The overall MATE process and terms were defined in section 2. The table below reflects how the author of the present thesis has interpreted each of those terms in the context of a food ingredient bioprocess:

Table 7 – MATE terms interpretation in bioprocess context

Term	Interpretation in the context of the system under study in this thesis
Mission Concept/Objective	Successful production of a biocompound. More specifically, lipase enzyme and DHA. In section 4.2.2 the mission objective will be properly stated as a system problem statement using the To-By-Using framework.
Systems	Lipase and DHA bioprocess
Attribute	Physicochemical characteristics of the compounds and overall characteristics of the bioprocess. For example, biocompound characteristics such as purity, stability, enzymatic activity, etc. and bioprocess characteristics such production yield, CO_2 emission, overall energy consumption, sustainability etc. It is important to point out that because 'attribute' includes process characteristics, it is a wider concept than product specifications.
Utility	A dimensionless parameter that reflects stakeholder satisfaction of an attribute. For example, for a biocompound purity of 60%, a numeric value between zero and one that reflects stakeholders overall satisfaction for a purity of 60%.
Mutli-Attribute Utility	A dimensionless parameter ranging from zero to one that reflects the stakeholder's satisfaction of the aggregation of all the attributes under consideration.
Design Variables	Bioprocess parameters that can be modified in order to improve the outcome of a (or several) attribute(s). Design variables are parameters such as fermentation pH, media, temperature, agitation, aeration, inoculum age, purification technique or specific column, etc.
Architecture	A given combination of design variables. For example, one architecture can be fermentation temperature 37°C, pH6, 150rpm and a second architecture can be fermentation temperature 40°C, pH7, 200rpm.
Tradespace	The set of all architectures under consideration

As a reminder to the reader, MATE steps introduced in section 2 are as follow:

1. Identify stakeholders
2. Define a mission objective/concept
3. Create a list of attributes
4. Determine design variables and map them to the attributes
5. Create a model that gives rise to utility curves
6. Evaluate architecture

In the following subsections each of these steps will be described in general terms for a bioprocess applicable to both case studies. However, when appropriate, important differences between the two case studies will be highlighted.

4.2.1. STAKEHOLDER AND STAKEHOLDER NEEDS

In order to define mission objective, first a stakeholder analysis was performed to identify needs the system (bioprocess) should fulfill. It was assumed for this analysis that the project was being carried out by a biotech startup developing a bioprocess to produce the food ingredient, in this case Lipase and/or DHA. The following stakeholder map shows the relationship between stakeholders. A full description of the stakeholders taken into consideration and their needs is presented in Appendix C.

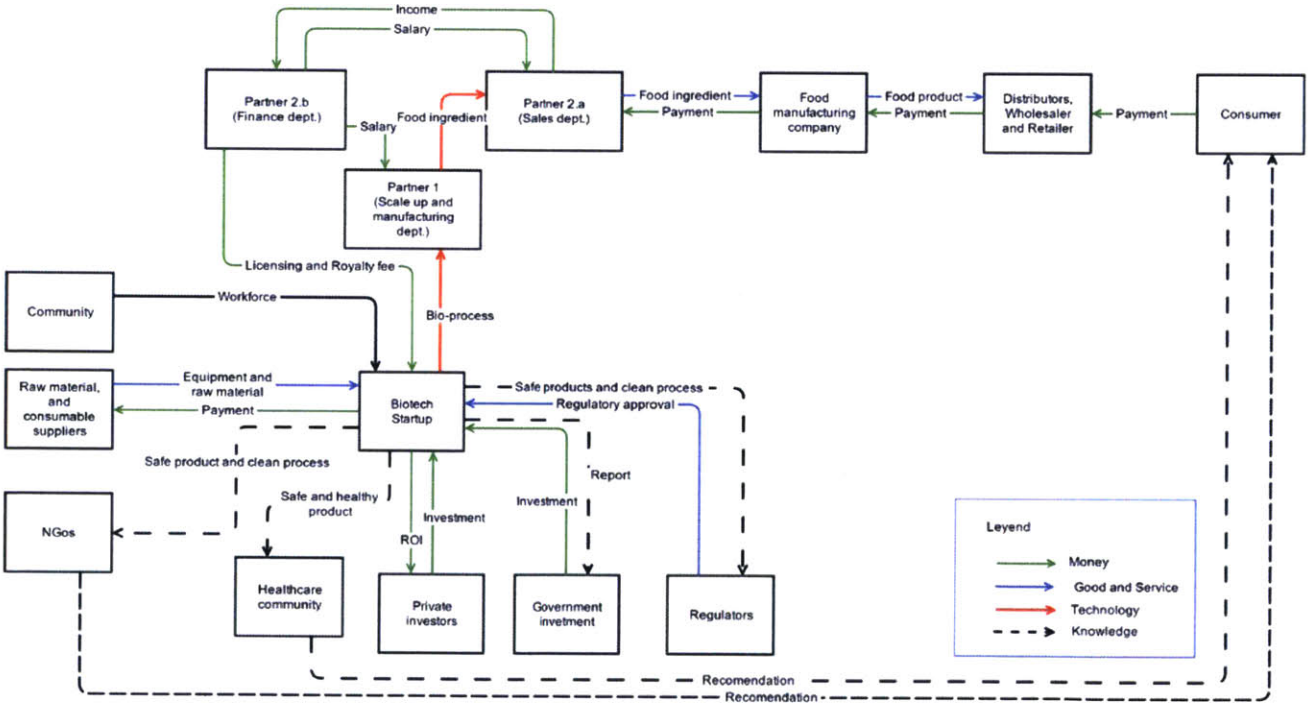


Figure 19 – Stakeholder map

For this analysis, it was assumed that the 'biotech startup' has a partnership agreement with a bigger ingredient manufacture company, where 'Partner 1'is the new product development and manufacturing department 'Partner 2.a' is the sales department and 'Partner 2.b' is the finance department. The characteristics of this partnership is as follow: a) biotech startup is in charge of the process development, b) partner 1 is in charge of the scaling up c) both find investment for each activity independently d) biotech startup does not pay the bigger company, as they are

partners in this project e) biotech startup receives licensing payment and royalty on sales from finance department.

Amongst the stakeholders presented in figure 19, some of them will pose requirement based on their need on process attributes and other on product attributes. For example, 'partner 1', and 'regulators' (EPA), would dictate process characteristics, while 'consumers', 'regulators' (FDA), and 'healthcare community' will influence the product attributes. From this analysis it was concluded that 'consumer' were the main drivers of product attributes in this system. Also, while the consumer does not buy ingredients directly, they are the main drivers for the monetary flow and influence other stakeholder's needs.

In a real-world setting, consumer needs would be determined based on surveys or focus groups. For this thesis, since the main objective is not to determine consumer needs, these were identified by the author based on experience and general sense base on online articles promoting DHA and lipase ingredients.

Table 8 – Ingredient market end-customer (consumer) needs

Attribute	Description
Healthy	Determined by the bio-compound decision/choice from the biotech startup and food manufacturing company, and promoted by healthcare community.
Tasty	Determined by the decision/choice from the biotech startup and food manufacturing formulation efforts.
Accessible	Provided by distribution channel, dependent on the ingredient and formulation stability developed by the biotech startup and the food manufacturing company.
Safe	Determined by the bioprocess developed by the biotech startup, ingredient manufactured by Partner 1, food product manufactured by Partner 2 and approval provided by regulator.
Affordable	Among other factor this need is fulfilled by a cost competitive bioprocess developed by the biotech startup.

4.2.2. MISSION OBJECTIVE STATED IN A TO-BY –USING FRAMEWORK

The To-By-Using framework was used to establish the system problem statement that describes the goal of the bioprocess system under study. The To-By-Using framework has the structure described below (Crawley, Cameron, & Selva, 2016):

To…[the statement of (solution-neutral functional) intent]
By verb-ing [statement (solution-specific) of function]
Using [the statement of form]

Figure 20 shows a detail view of the To-By-Using framework and a schematic view using Object Process Methodology (OPM)

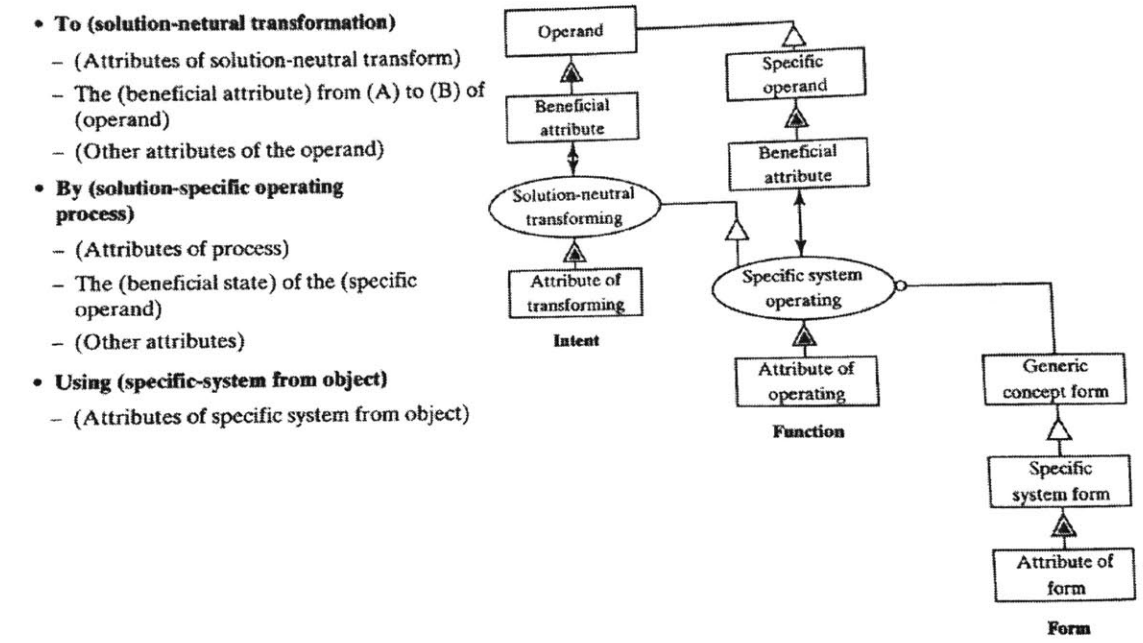


Figure 20 – To-By-Using framework for formulating System Problem Statement and graphical representation in OPM.

OPM was developed by Professor Dov Dori at Technion with the goal of unifying the object-and process- oriented paradigms for describing systems in a single methodology. Rectangle represent object. Circles represent process. A black triangle inside another represents a characterization link, a white triangle represents subclass (specification). An extensive description of the OPL (Object Process Language) symbols and meaning is provided in Appendix D. Source: (Crawley, Cameron, & Selva, 2016)

In systems engineering 'form' refers to what the system 'is', while 'function' refers to what the system 'does'. A solution neutral function is "the function of a system stated without reference to how it is achieved" (Crawley, Cameron, & Selva, 2016). Thus, the first step was to define the intent without expressing how the system would achieve it. This is done by focusing on the value delivered to the primary beneficiary needs. Then, "by" states the solution specific function. This is, states how the intent will be achieved. Finally, "using" states the form or what will perform the function. The system problem statement developed for the 'Food Ingredient Bioprocess System' is as follow:

To improve quality of food
By manufacturing healthy, tasty, safe, accessible and affordable food ingredients
Using cost efficient and environmentally friendly bioprocesses.

The schematic view for this problem statement using OPM is presented below:

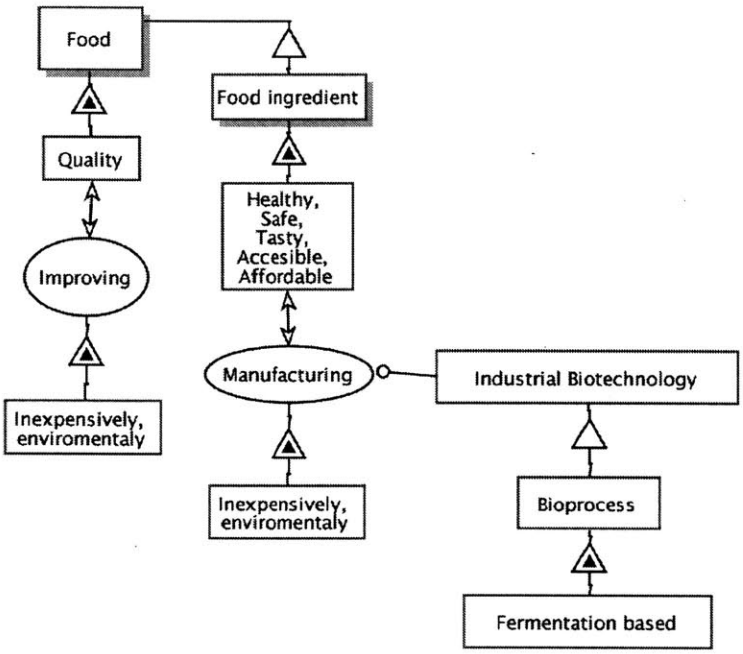


Figure 21 – OPM representation of System Problem Statement for a 'Food Ingredient Bioprocess System'

4.2.3. ATTRIBUTE-VARIABLE MAPPING USING DESIGN VALUE MATRIX (DVM)

I have so far identified the consumer as the primary beneficiary stakeholder, described its relationship with other stakeholders through a stakeholder map, define its needs and stated a system problem statement using the To-By-Using framework. The next step of MATE is to create a list of attributes and map them to design variables.

As explained in section 2, 'Attribute' is traditionally defined as 'a decision-maker perceived metric that measures or determines how well the defined objective is met'. For example, in the case of DHA, an attribute that might be perceived by the decision-maker that determines how well the need of "tasty" is met is perhaps "no fishy flavor". However, this attribute is very difficult to measure. Specially at the begging of the design process, when the physical product is not available. Thus, in this thesis, 'attribute' will be interpreted as measurable product or process characteristics that would lead to the satisfaction of the primary stakeholder's needs. Thus, the following table describes how each need was interpreted and the chosen attribute(s) used for the rest of the analysis.

Table 9 – DHA attribute table

Need	Interpretation	Attribute
Healthy	DHA is perceived as healthy mainly due to the effective communication of its health benefits. However, rancid oil is unhealthy and also unsafe.	Low oxidation level
Tasty	Off flavor in Omega-3 products can be produced by oxidation. It is usually prevented and or decreased by deodorizing step during the purification process, microencapsulation and masking with flavoring systems during the formulation stage.	Low oxidation level
Safe	In order to be considered safe, the new DHA ingredient produced by specific strain of microorganisms should have GRAS (Generally Recognized A Safe) status reviewed by the FDA. Also, if its intended use is for infant formula, it must be free of EPA. Finally, it has to be free of toxins and heavy metals.	High purity (low or absence of heavy metals, toxins and EPA) GRAS status
Accessible	It order for it to be accessible, it must be easily stored and transported. Thus it needs to be stable and not easily (or less easily) oxidized. This is usually achieved by adding antioxidants to the formulation.	Stable oxidation level in time and/or at different temperatures
Affordable	In order for it to be affordable, the process needs to be cost efficient. Thus, productivity and/or yield needs to be high.	High productivity and/or yield

Table 10 – Lipase attribute table

Need	Interpretation	Attribute
Healthy and tasty	While lipase is sold also as a supplement, in the food and beverage segment its mainly used as a catalyst and not as a functional food ingredient. Thus, the perceived 'healthy' and 'tasty' qualities does not come from the lipase itself, but rather the food product of which manufacturing lipase take part of.	Not applicable
Safe	In order to be considered safe the enzyme preparation manufactured by a specific microorganisms should have GRAS status reviewed by the FDA.	GRAS status
Accessible	In this case, as the enzyme is used in food preparation, consumers does not have access to the enzyme, but rather the food product prepared using the enzyme.	Not applicable
Affordable	In order for it to be affordable, enzyme manufacturers need to have access to big quantities of highly active enzyme. Thus the enzyme preparation has to have high activity, productivity and stability.	High productivity and or yield High enzymatic activity Stable enzymatic activity through time and/or different temperatures

It is important to point out that the objective of this thesis in centered on the use of DoE in combination with the MATE method, which will be further explained in the next section. Thus, this list of attributes is not by any means complete nor extensive. It was developed with the

objective of illustrating and translating this method used in other industries into the development of a bio-compound. A survey and deeper analysis using methods such as QFD (Quality Function Deployment) would be required for a complete list of attributes, which is outside the scope of the present thesis.

Once the attributes were selected, an attempt to map these attributes to design variables was made. However, it was evident that unlike aerospace systems, biological systems are much more complex, in that a given attribute cannot be mapped to one or to a group of design variables. For example, for a Space Based Radar (SBR) used to take images of stationary ground targets, the design variables 'orbital altitude' and 'radar aperture size' would be mapped to the attribute 'high resolution image' (Spaulding, 2003). However, for a DHA produced through bioprocessing of an algal strain, it is unclear what design variable has an effect on the attribute "low oxidation level". On the other hand, potentially all design variables can affect the attribute "High productivity/yield". The matrix bellow, usually referred to as DVM (Design-Value Matrix) graphically shows this fact.

	Process Steps	Upstream						Fermentation							Downstream																							Formulatio and others
	Activities (unit operation) within each step	Strain engineering	Inoculum (seed) preparation					Fermentation							Cell Harvest (centrifugation)		Extraction and saponification							Methylation				Winterization			Urea complexation							Formulatio and others
	Design variables within unit operation	inoculum strain	inoculation medium	inoculum growth temperature	inoculum gowth pH	inoculum growth agitation	inoculum growth time	Fermentation medium	Fermentation temperature	Fermentation pH	Fermentation agitation	Fermentation aeration	Fermentation time	Fermentor design	Centrifugation time	Centrifugation RCF	Ethanol concentration	KOH concentration	Temperature	Agitation (RPM)	Time	HCl concentration	Hexane concentration	Methanol concentration	Sulfuric acid concentration	Temperaure	Time	Hexane concentration	Temperature	Time	Urea concentration	temperature	Stirring rate	Centrigufation RPM				
Attributes	Oxidation Level								?			?						?	?	?		?	?	?	?	?		?	?		?	?					x
	Purity																x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x			
	Productivity/yield	x	x	x	x	x	x	x	x	x	x	x	x	x			x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x			
	GRAS status																																					x

Figure 22 – DHA DVM (Design-Value Matrix)

4.2.4. APPLICATION OF DOE TO MODEL UTILITY CURVES

Based on the fact that in bioprocess systems it is very difficult to determine/map a given attribute to the corresponding design variable(s) (figure 22), the use of a tool to discriminate which are the main design variables that have the most effect on a given attribute is necessary. Once the main design variables are determined a mathematical model describing how the selected design

variables affect the given attribute, referred in this thesis as "attribute-curve" or "attribute-function", is required. The present thesis proposes the use of two DoE designs commonly used in bioprocesses to accomplish this. Namely, Plackett-Burman (PB) designs (Plackett & Burman, 1946) to down select the design variables and Response Surface Methodology (RSM) to produce attribute-curves. As explained in section 2.2, Design of experiments deals with quantifying how process inputs affect process output using a minimum number of runs. Thus, it is useful to create a model between design variables and attribute. The resulting attribute-curve can then be converted into SUFs using a linear function based on expert opinion of what the maximum (x_{imax}) and minimum (x_{imin}) acceptable values for attribute i should be. In this way,

$$SUF_i(x_{imax}) = u_i(x_{imax}) = 1 \text{ and } SUF_i(x_{imin}) = u_i(x_{imin}) = 0.$$

The following figure schematically explains the proposed MATE incorporating PB design and RSM to create utility curves.

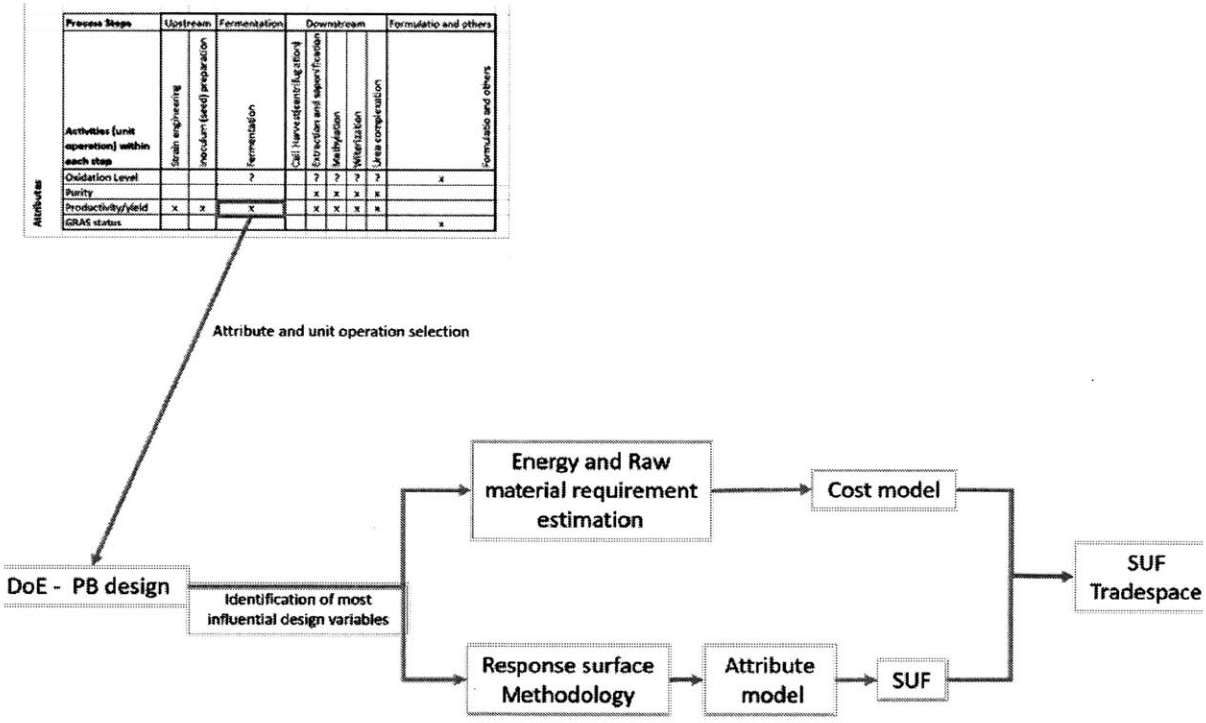


Figure 23 – Single utility function –SUF- obtained through DoE

The first step is to map attributes to unit operations instead of design variables, as the relevant design variables will be revealed after a PB design experiment. The down selection of design variables is performed per unit operation, per attribute. Once the relevant design variables are identified, a RSM design is performed in order to develop attribute models. The attribute model or function is then translated into single utility functions (SUF), which can be plotted in a trade space showing single utility curves. Finally, single utility curves can be aggregated into a multi-utility curve by giving it a weight-based on consumer preference, as shown in the schematic diagram bellow:

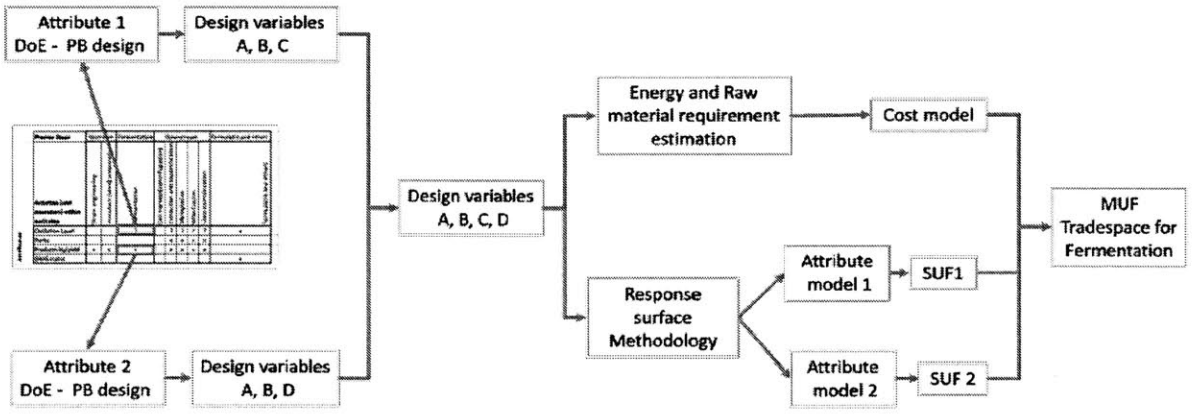


Figure 24 – Multi-utility-function –MUF- obtained through DoE

In the example above (figure 24), unit operation 'Fermentation' might have an effect on two attributes: 'attribute 1' and 'attribute 2'. A PB design experiment is performed and design variables 'A', 'B' and 'C' were found to be relevant for 'attribute 1', whereas design variables 'A', 'B' and 'D' were relevant for 'attribute 2'. In the next step a RSM is performed taking into account all design variables, including those that are relevant only for one attribute. As a result of the RSM, single utility functions can be developed and plotted. Finally, a multi-attribute utility function is developed based on the weight given for each attribute according to customer preference. As a result, a multi-attribute Tradespace plot for unit operation 'Fermentation' can created as a decision making tool.

In the following subsections, this process is exemplified in the two case studies mentioned earlier, lipase and DHA production.

4.3. CASE STUDY 1: ALGAL DHA PRODUCTION

DHA case study was based on Song et al. (2007) study on "Optimization of fermentation parameters for the biomass and DHA production of *Schizochytrium limacinum* OUC88 using response surface methodology", published in the journal Process Biochemistry. While both biomass productivity and DHA productivity were measured in the paper, only the product, DHA, was considered in this analysis. As shown in table 9, the attribute 'DHA productivity' would be correlated with customer need of 'affordable'. Song et al. performs a PB design in his paper to identify the factor (or 'design variables' as we will refer to them from now on) that has the greatest effect on DHA production. This case study was chosen to exemplify how PB design and RSM can potentially be integrated in the MATE method to obtain SUFs.

4.3.1. ATTRIBUTE MODEL

Song et al. analyzed 10 factors in their PB design: temperature (T, C°), aeration rate (Q, volume of air per volume of medium per minute VVM), pH, agitation (R, rpm), inoculum volume (I, %), fermentation volume (V, L), fermentation pressure (P, Mpa) inoculum age (IA), harvesting time (HT, h), and Tween 80 concentration (Tw, mL). The table below was obtained from the original paper and shows the high (+1) and low (-1) levels for each parameter.

Table 11 – Case Study 1: DHA PB design variables (factors) range

Source: (Song X. , Zhang, Kuang, Zhu, & Guo, 2007)

Range of different factors studied in the Plackett–Burman design

Variable	Variable code	Low level (−1)	High level (+1)
Temperature/T (□)	X_1	23	26
Aeration rate/Q (L min^{-1} L^{-1})	X_2	1.02	1.48
pH	X_3	6	7
Agitation/R (rpm)	X_4	150	250
Inoculum volume/I (%)	X_5	7	10
Fermentation volume/V (L)	X_6	6	8
Fermentation pressure/P (Mpa)	X_7	0.06	0.08
Inoculum age/IA	X_8	Mid-exponential phase	Stationary phase
Harvesting time/HT (h)	X_9	108	132
Tween 80 concentration/Tw (ml)	X_{10}	2	10

The following table summarizes the results Song et al obtained for the PB design, where the variables with confidence level greater than 95%, thus influence DHA production are: 1)

temperature (T, C°), 2) aeration rate (Q, volume of air per volume of medium per minute VVM) and 3) agitation (R, rpm).

Table 12 – Case Study 1: DHA PB design results

Source: adapted from (Song X. , Zhang, Kuang, Zhu, & Guo, 2007)

Run#	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	DHA (g/L)
1	1	-1	1	-1	-1	-1	1	1	1	-1	1.7
2	1	1	-1	1	-1	-1	-1	1	1	1	3.1
3	-1	1	1	-1	1	-1	-1	-1	1	1	3.9
4	1	-1	1	1	-1	1	-1	-1	-1	1	2.4
5	1	1	-1	1	1	-1	1	-1	-1	-1	3.8
6	1	1	1	-1	1	1	-1	1	-1	-1	2.5
7	-1	1	1	1	-1	1	1	-1	1	-1	4.7
8	-1	-1	1	1	1	-1	1	1	-1	1	3.9
9	-1	-1	-1	1	1	1	-1	1	1	-1	3.5
10	1	-1	-1	-1	1	1	1	-1	1	1	2
11	-1	1	-1	-1	-1	1	1	1	-1	1	3.5
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	3
Effect	**-1.167**	**0.807**	0.057	**0.810**	0.183	-0.160	0.203	-0.260	-0.067	-0.053	
S.E	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	
t	-20.588	14.235	1.000	14.294	3.235	-2.824	3.588	-4.588	-1.176	-0.941	
Pr>|t|	0.031	0.045	0.500	0.044	0.191	0.217	0.173	0.137	0.448	0.519	

These three design variables were used in the central composite RSM design. The RSM design and DHA yield was obtained from Song et al. Based on this information, a second-order polynomial equation was found to best describe the attribute DHA yield (g/L), as a function of the design variables temperature, aeration and agitation.

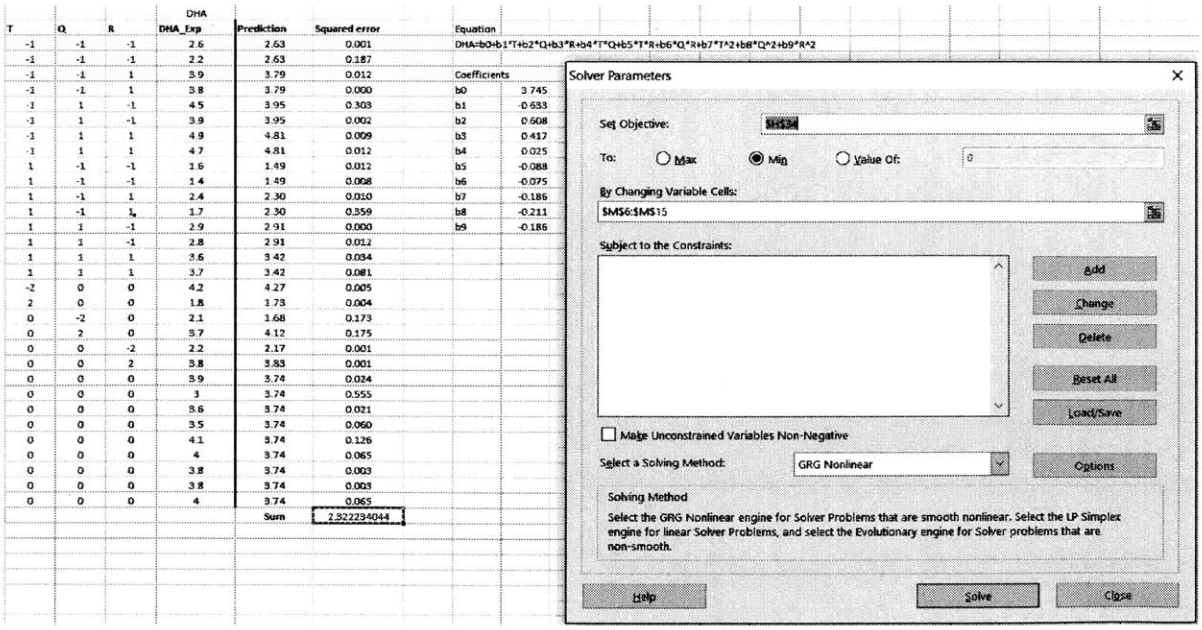


Figure 25 – Case Study 1: DHA attribute function

The equation obtained was:

$$Y_{DHA} = 3.745 - 0.633T + 0.608Q + 0.417R + 0.025TQ - 0.088TR - 0.075\text{QR} - 0.186T^2 - 0.211Q^2 - 0.186R^2$$

4.3.2. SUF MODEL

From literature, the maximum published productivity of DHA by *Schizochytrium sp.* found was 10-12g/L*day (Martek patent, US7732170). This data was used as the maximum attribute value (x_{imax}). In a real-life scenario, the minimum acceptable value for the attribute productivity would be established based on expert's opinion. However, in this case study, for simplicity the minimum acceptable value for this attribute (x_{imin}) was equated to the minimum experimental yield obtained (1.49 g/L). On the other hand, through PB design, Song et al (2007), proved that harvesting time, with 108 hrs in the low level (-1), has no effect on DHA yield. Therefore, it was assumed that for the RSM experiment the harvesting time used was 108hrs (4.5 days). Thus, x_{imin} = 0.33 g/L*day. Based on these information, the following two equations were used to develop a Single-Utility Function (SUF) for the attribute 'productivity':

$$1 = 12\,m + b$$

$$0 = 0.33m + b$$

The solution for this simple linear $u_i(x_{ij})$ system yields the utility function that transforms the attribute rating into a utility value between 0 and 1:

$$SUF_{productivity}(x_{productivityj}) = 0.086(x_{productivityj}) - 0.028$$

Where $SUF_{productivity}$ is single utility function for attribute *'productivity'*, $x_{productivityj}$ is the attribute rating (the raw score) for alternative *j* of the attribute *'productivity'*.

In order to convert the previously obtained utility model to SUF model the predicted DHA yield (g/L) was first translated into productivity (g/L*day) by dividing the yield by 4.5 hrs. Then using the equation above the corresponding utility scores were obtained.

Table 13 – Case Study 1: DHA SUF (Single utility function)

RUN#	Predicted DHA yield	Predicted DHA productivity	Utility score	Utility function	
	g/L	g/L*day		SUF =m (predicted DHA productivity)+b	
1	2.63	0.59	0.022		
2	2.63	0.59	0.022	m	0.08569
3	3.79	0.84	0.044	b	-0.0283
4	3.79	0.84	0.044		
5	3.95	0.88	0.047		
6	3.95	0.88	0.047		
7	4.81	1.07	0.063		
8	4.81	1.07	0.063		
9	1.49	0.33	0.000		
10	1.49	0.33	0.000		
11	2.30	0.51	0.016		
12	2.30	0.51	0.016		
13	2.91	0.65	0.027		
14	2.91	0.65	0.027		
15	3.42	0.76	0.037		
16	3.42	0.76	0.037		
17	4.27	0.95	0.053		
18	1.73	0.39	0.005		
19	1.68	0.37	0.004		
20	4.12	0.92	0.050		
21	2.17	0.48	0.013		
22	3.83	0.85	0.045		
23	3.74	0.83	0.043		
24	3.74	0.83	0.043		
25	3.74	0.83	0.043		
26	3.74	0.83	0.043		
27	3.74	0.83	0.043		
28	3.74	0.83	0.043		
29	3.74	0.83	0.043		
30	3.74	0.83	0.043		
31	3.74	0.83	0.043		

This utility score informs the decision maker about how far the attribute 'productivity' is from the ideal value (x_{imax}). Also, given that now it is a utility score, it can be added to other attribute scores as will be exemplified in the second case study.

4.3.3. COST MODEL

The second variable in the Tradespace is cost. The objective is to be able to differentiate what architecture has the highest cost. For this, a differential cost, rather than a total fermentation process cost, is enough. Assuming all other design variables (besides from temperature, aeration

and agitation) to be constant, the differential cost will be defined by the cost of heating up and cooling down the bioreactor, the cost of injecting oxygen into the reactor and the cost of agitation. From amongst these three variables, heating/cooling and agitation cost are more significant. However, since the cost of heating and cooling depends on the ambient temperature (and therefore on the season and the region where facility is located) the power consumption for stirring was used as an approximation for cost.

The power density per impeller was calculated using the following formula:

$$P = P_0 \rho N^3 D_i^5$$

Where P is the power consumption per impeller, P_0 the power number (5.5) for a given type of impeller, N rotation per second (rps) and Di is the diameter of the impeller (0.076m). The resulting power consumption was multiplied by the number of impellers, in this case 3, and divided by the volume (Vz), in this case 7L.

Based on this information, the power consumption per architecture was calculated and shown in the following table:

Table 14 – Case Study 1: DHA Power density calculation

RUN#	RPM	N	P	P/Vz
	1/min	1/seg	W	W/L
1	150	2.500	0.658	0.09
2	150	2.500	0.658	0.09
3	250	4.167	3.046	0.43
4	250	4.167	3.046	0.43
5	150	2.500	0.658	0.09
6	150	2.500	0.658	0.09
7	250	4.167	3.046	0.43
8	250	4.167	3.046	0.43
9	150	2.500	0.658	0.09
10	150	2.500	0.658	0.09
11	250	4.167	3.046	0.43
12	250	4.167	3.046	0.43
13	150	2.500	0.658	0.09
14	150	2.500	0.658	0.09
15	250	4.167	3.046	0.43
16	250	4.167	3.046	0.43
17	200	3.333	1.560	0.22
18	200	3.333	1.560	0.22
19	200	3.333	1.560	0.22
20	200	3.333	1.560	0.22
21	100	1.667	0.195	0.03
22	300	5.000	5.264	0.74
23	200	3.333	1.560	0.22
24	200	3.333	1.560	0.22
25	200	3.333	1.560	0.22
26	200	3.333	1.560	0.22
27	200	3.333	1.560	0.22
28	200	3.333	1.560	0.22
29	200	3.333	1.560	0.22
30	200	3.333	1.560	0.22
31	200	3.333	1.560	0.22

The average industrial electricity rate in Massachusetts is 12.57¢/kWh (Electricity Local, 2016). Assuming a production of $50m^3$ the estimated cost is:

Table 15 – Case Study 1: DHA power consumption cost calculation

RUN#	P/Vz	¢/L	$/50M3		
	W/L			Harvest time (hr)	108
1	0.09	0.13	62.88	Electricity cost ($/kWh)	0.1257
2	0.09	0.13	62.88	Electricity cost ($/Wh)	0.000126
3	0.43	0.58	291.67		
4	0.43	0.58	291.67		
5	0.09	0.13	62.74		
6	0.09	0.13	62.74		
7	0.43	0.58	291.31		
8	0.43	0.58	291.31		
9	0.09	0.13	62.88		
10	0.09	0.13	62.88		
11	0.43	0.58	291.66		
12	0.43	0.58	291.66		
13	0.09	0.13	62.74		
14	0.09	0.13	62.74		
15	0.43	0.58	291.30		
16	0.43	0.58	291.30		
17	0.22	0.30	149.12		
18	0.22	0.30	149.11		
19	0.22	0.30	149.35		
20	0.22	0.30	148.86		
21	0.03	0.04	18.55		
22	0.74	1.01	503.97		
23	0.22	0.30	149.11		
24	0.22	0.30	149.11		
25	0.22	0.30	149.11		
26	0.22	0.30	149.11		
27	0.22	0.30	149.11		
28	0.22	0.30	149.11		
29	0.22	0.30	149.11		
30	0.22	0.30	149.11		
31	0.22	0.30	149.11		

4.3.4. TRADESPACE

With the information obtained in the previous sections, a SUF Tradespace for the attribute 'productivity' was plotted.

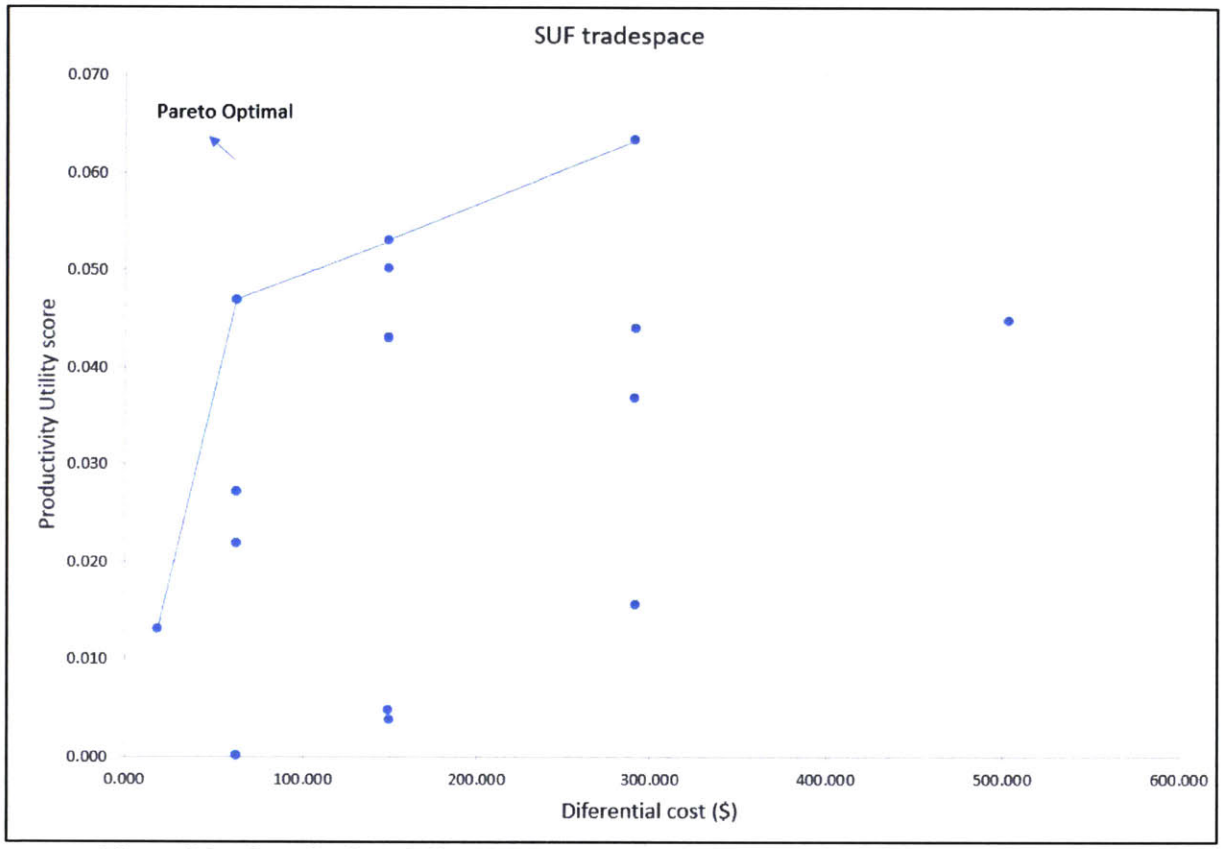


Figure 26 – Case Study 1: DHA SUF Tradespace

The Pareto optimal in this case is the left upper corner, where utility score in the highest and the differential cost is the lowest.

The benefit of having utility scores, as opposed to productivity, is that it allows the addition of different attributes, with different units into one parameter. This will be exemplified in case study 2. In this case study, however, since only one attribute is being analyzed, the following plot of productivity v/s differential cost will facilitate the analysis.

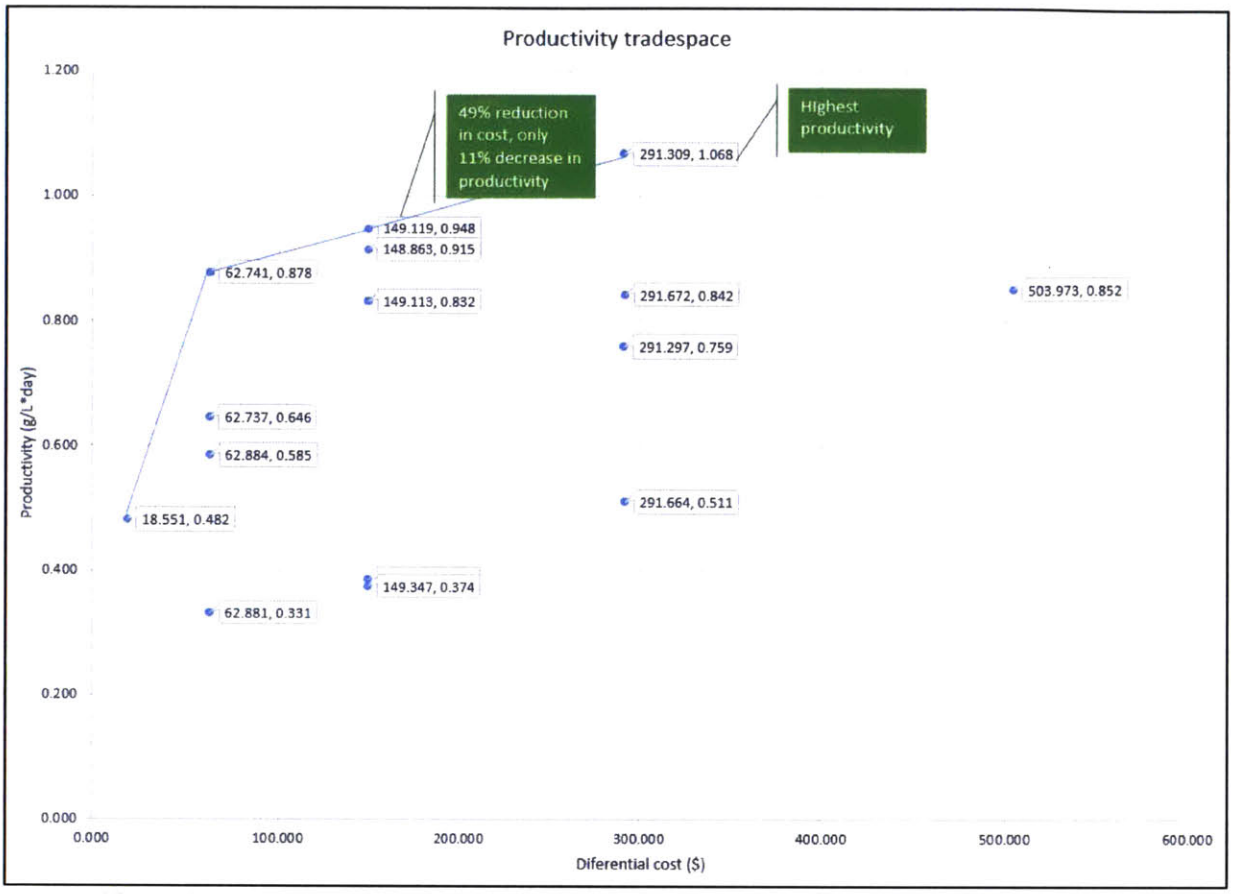


Figure 27 – Case Study 1: DHA Productivity Tradespace

In a traditional optimization process, where cost is not analyzed in parallel, the optimal architecture would be represented by point (291.309, 1.068) in the plot. This architecture, with the following design variables: 23°C, 1.48 L/min*L and 250 rpm results in an optimal productivity of 1.068g/L*day of DHA. This corresponds to a cost of $291.309 higher than the cheapest architecture analyzed. However, if we follow the Pareto frontier, moving towards the Pareto optimal, the following non-dominated architecture is represented by the point (149.119, 0.948). This architecture has the following design variables: 21.5°C, 1.25 L/min*L and 200 rpm. While the resulting productivity in this case is 11% lower (0.948g/L*day compared to 1.068g/L*day), the differential cost is 49% lower ($149 compared to $291). Whether this architecture is a better option or not will depend on the price of DHA. In other words, does the loss of profit associated with selling an additional 0.12g/L*day of DHA justify saving $142 per run? The price of DHA depends not only on the end market segment (food and beverage, infant formula, supplement, animal feed, clinical nutrition, etc.), but also on other attributes, such as purity. Thus the detail

analysis and construction of a scenario to thoroughly analyze what the best architecture is, falls beyond the scope of the present thesis. However, what can be concluded is that taking into consideration fermentation unit operation in isolation, the productivity optimal alone might not yield enough information to make the best decision. Furthermore, Tradespace might be a tool that can support decision making when several objectives needs to be met.

4.4. CASE STUDY 2: HYPERTHERMOSTABLE LIPASE FERMENTATION

The Lipase case study was based on Rathi et al (2002) paper "Statistical medium optimization and production of a hyperthermostable lipase from *Burkholderia cepia* in a bioreactor" published in Journal of Applied Microbiology. In this paper, design variables down-selection was not done through a PB design, but rather a 'one-factor-at-a-time' method published in an earlier paper (Rathi, Saxena, & Gupta, 2001). However, two attributes (or output variables) are measured in the RSM, namely yield and enzyme specific activity. Thus, this paper was chosen to exemplify how a multi-attribute-utility can be obtained from two SUFs.

4.4.1. ATTRIBUTE MODEL

According to Rathi et al. (2001) 'one-factor- at-a-time' analysis, the relevant design variables for lipase attributes 'production yield' (U/mL) and specific activity (U/mg) are: glucose concentration, palm oil concentration, incubation time, inoculum age and agitation. The range and the codes for these 5 design variables are shown in the table below:

Table 16 – Case Study 2: Lipase design variables range

Source: (Rathi, Goswami, Sahai, & Gupta, 2002)

	Range of levels					
Variables	Actual	Coded	Actual	Coded	Actual	Coded
Glucose (mg ml^{-1}): A	2	−1	8	0	14	+1
Palm oil (% v/v): B	0	−1	1	0	2	+1
Incubation time (h): C	20	−1	40	0	60	+1
Inoculum density (%): D	1	−1	2	0	3	+1
Agitation (rev min^{-1}): E	100	−1	200	0	300	+1

These five design variables were used in a RSM central composite faced center design to develop a attribute model for 'yield' and 'specific activity'. The experimental design and the experiment results were obtained from Rathi (2002). Similar to case study 1, these results were used to fit a second-order polynomial equation.

Table 17 – Case Study 2: Lipase yield attribute function

							lipase yield U/mL				
2	RUN#	Glu	Palm oil	Incubation time	Inoculum density	Agitation	Observed	Prediction	Squared error		
3	1	0	1	0	0	0	8.2	5.88	5.404		
4	2	0	0	0	1	0	28	24.71	10.841		
5	3	0	-1	0	0	0	3.8	3.09	0.511	Coefficients	
6	4	1	-1	1	1	-1	0.5	0.44	0.004	b0	11.87617
7	5	1	1	1	1	1	8.1	8.92	0.673	b1	0.354867
8	6	0	0	0	0	0	9.3	11.88	6.637	b2	1.395127
9	7	1	-1	-1	-1	-1	8.3	6.57	2.997	b3	-1.48402
10	8	0	0	0	0	0	12	11.88	0.015	b4	0.877092
11	9	1	1	-1	-1	1	12.1	12.21	0.012	b5	1.905691
12	10	1	1	1	-1	-1	6.6	5.82	0.614	b6	0.542979
13	11	-1	-1	1	-1	-1	2.6	2.67	0.004	b7	-1.25548
14	12	-1	-1	-1	1	-1	6.7	6.46	0.056	b8	-0.49299
15	13	0	0	1	0	0	11.5	7.00	20.283	b9	-0.31267
16	14	-1	0	0	0	0	15.3	9.58	32.770	b10	1.230482
17	15	1	0	0	0	0	7.6	10.29	7.210	b11	0.667979
18	16	0	0	-1	0	0	8.5	9.96	2.144	b12	0.287653
19	17	0	0	0	0	0	9.6	11.88	5.181	b13	-0.65548
20	18	0	0	0	0	0	8	11.88	15.025	b14	-0.72516
21	19	0	0	0	0	0	9.7	11.88	4.736	b15	-0.46267
22	20	1	1	-1	1	-1	11.5	10.41	1.181	b16	-1.94581
23	21	-1	1	-1	-1	1	8.5	7.54	0.917	b17	-7.39586
24	22	1	-1	-1	1	1	11.5	11.37	0.016	b18	-3.39582
25	23	0	0	0	0	-1	2.8	5.92	9.764	b19	11.9541
26	24	-1	1	1	1	-1	7.9	8.61	0.505	b20	-4.04582
27	25	-1	-1	-1	-1	1	8.1	9.06	0.921		
28	26	1	-1	1	-1	1	3.3	3.48	0.031		
29	27	0	0	0	0	1	15.9	9.74	37.994		
30	28	-1	1	1	-1	1	7.5	9.41	3.637		
31	29	-1	1	-1	1	1	10.4	12.00	2.574		
32	30	-1	-1	1	1	1	3.5	5.17	2.790		
33	31	0	0	0	-1	0	22.7	22.95	0.064		
34		0	0	0	0	0	10.5	11.88	1.894		
35								Sum	177.406		

Table 18 – Case Study 2: Lipase activity attribute function

						lipase specific activty U/mg				
RUN#	Glu	Palm oil	Incubation time	Inoculum density	Agitation	Observed	Prediction	Squared error		
1	0	1	0	0	0	51	41.08	98.335		
2	0	0	0	1	0	96	88.38	58.090		
3	0	-1	0	0	0	20	23.38	11.413	Coefficients	
4	1	-1	1	1	-1	13	7.16	34.127	b0	60.27
5	1	1	1	1	1	48	50.16	4.672	b1	0.31
6	0	0	0	0	0	57	60.27	10.726	b2	8.85
7	1	-1	-1	-1	-1	53	48.64	18.970	b3	-11.49
8	0	0	0	0	0	55	60.27	27.826	b4	0.15
9	1	1	-1	-1	1	58	60.38	5.656	b5	4.17
10	1	1	1	-1	-1	34	29.37	21.468	b6	-0.65
11	-1	-1	1	-1	-1	14	11.03	8.801	b7	-2.10
12	-1	-1	-1	1	-1	34	34.06	0.004	b8	0.52
13	0	0	1	0	0	54.5	58.99	20.157	b9	1.31
14	-1	0	0	0	0	67	50.42	274.992	b10	4.35
15	1	0	0	0	0	41	51.05	100.906	b11	1.48
16	0	0	-1	0	0	93	81.97	121.602	b12	-1.06
17	0	0	0	0	0	57	60.27	10.726	b13	3.02
18	0	0	0	0	0	55	60.27	27.826	b14	0.81
19	0	0	0	0	0	56	60.27	18.276	b15	1.18
20	1	1	-1	1	-1	53	51.39	2.578	b16	-9.54
21	-1	1	-1	-1	1	54	55.27	1.612	b17	-28.04
22	1	-1	-1	1	1	50	52.44	5.951	b18	10.21
23	0	0	0	0	-1	15	34.56	382.519	b19	27.96
24	-1	1	1	1	-1	44	43.78	0.047	b20	-21.54
25	-1	-1	-1	-1	1	46	50.04	16.361		
26	1	-1	1	-1	1	18	17.41	0.346		
27	0	0	0	0	1	69	42.90	681.002		
28	-1	1	1	-1	1	37	40.77	14.191		
29	-1	1	-1	1	1	47	53.79	46.171		
30	-1	-1	1	1	1	19	22.83	14.656		
31	0	0	0	-1	0	87	88.08	1.174		
	0	0	0	0	0	55.5	60.27	22.801		
							Sum	2063.977		

4.4.2. SUF MODEL

Information about the 'yield' and 'specific activity' values, before purification for the production of industrial hyperstable alkaline lipase, was not found. Thus, for this exercise, the values reported by Bhosale et al (2016) were used as benchmark. The table below shows the results obtained by Bhosale et al.

Table 19 – Case Study 2: Lipase attribute benchmark

Source: (Bhosale, Shaheen, & Kadam, 2016)

	Protein content (mg/mL)	Total activity	Specific activity (U/mg)	Purification fold	Yield (%)
Crude	0.290	153990	177	1	100
Ammonium sulphate precipitation	0.252	13275	351	1.98	8.62
Dialysis	0.143	5901.5	825	4.67	3.83
DEAE-cellulose column	0.355	3055.96	2152.08	12.15	1.98

On table 19, the line 'crude' refers to crude extract, which is centrifuged broth after 4 days of incubation, thus cell-free supernatant. This is the same step in which Rathi et al (2001) performed measurements of yield and specific activity. In both cases the microorganism were grown in shake flasks. However, in Bhosale et al. microorganisms were harvested after 4 days, whereas Rathi tested different incubation times. In order to account for this difference of incubation time, productivity was calculated by multiplying specific activity (U/mg) by protein content (mg/mL) and divided by 4 days (96 hours) , thus 177 U/mg * 0.290 mg/mL * 1/96 hours = 0.53 U/mL*h. Since this is productivity in shake flask, it is expected that productivity in bioreactor will improve, due to superior aeration conditions in fermenters. However, bioreactor fermentation results was not reported for the complete set of architectures analyzed. Thus, shake flask productivity was used as a proxy instead, assuming 0.53U/mL*h as an average acceptable productivity (utility value of 0.5) and a maximum productivity set at twice this value. Therefore, the corresponding equations are:

$$x_{productivity\_0.5} = 0.53\frac{U}{mL * h} = 0.5$$

$$x_{productivity\_max} = 1.06\frac{U}{mL * h} = 1$$

$$SUF_{productivity}(x_{productivityj}) = 1.06(x_{productivityj}) - 0.1236$$

Where $SUF_{productivity}$ is the single utility function for attribute *'productivity'*, $x_{productivityj}$ is the attribute rating (the raw score) for alternative *j* of the attribute *'productivity'*. The 'specific activity' value used as benchmark was the value reported by Bhosale, 177 U/mg. In this case, this value was used as minimum acceptable value. A lower specific productivity suggests that from the total protein produced, a lower percentage corresponds to the lipase enzyme or that the lipase activity is much lower. Rathi, should be able to produce at least as much lipase activity per mg of total protein as Bhosale. The maximum value was set at 3 times the minimum value,

thus 531U/mg. With this maximum value of specific activity in crude extract, a specific activity of ~5,000 U/mg is expected for purified enzyme, which corresponds with the activity reported for Novozymes ® CAL-B lipase. The corresponding equations are:

$$x_{activity\_min} = 177\frac{U}{mg} = 0$$

$$x_{activity\_max} = 531\frac{U}{mg} = 1$$

$$SUF_{activity}(x_{activityj}) = 0.00282(x_{productivityj}) - 0.5$$

Where $SUF_{activity}$ is the single utility function for attribute *'specific activity'*, $x_{activityj}$ is the attribute rating (the raw score) for alternative j of the attribute '*specific activity*'.

The following table shows the utility scores for the attributes 'productivity' and 'specific activity'.

Table 20 – Case Study 2: Lipase SUF (Single-Utility Function)

RUN#	Yield (U/mL)	Incubation time (h)	Productivity (U/mL*h)	Specific activity (U/mg)	Productivity utility score	Specific activity utility score		
1	5.88	40.00	0.15	41.08	0.032	-0.384	**Productivity Utility function**	
2	24.71	40.00	0.62	88.38	0.531	-0.250	SUF =m (predicted Lipase productivity)+b	
3	3.09	40.00	0.08	23.38	-0.042	-0.434	m	1.0600
4	0.44	60.00	0.01	7.16	-0.116	-0.480	b	-0.1236
5	8.92	60.00	0.15	50.16	0.034	-0.358		
6	11.88	40.00	0.30	60.27	0.191	-0.330	**Specific activity Utility function**	
7	6.57	20.00	0.33	48.64	0.225	-0.363	SUF =m (predicted Lipase specific activity)+b	
8	11.88	40.00	0.30	60.27	0.191	-0.330	m	0.00282
9	12.21	20.00	0.61	60.38	0.524	-0.329	b	-0.5
10	5.82	60.00	0.10	29.37	-0.021	-0.417		
11	2.67	60.00	0.04	11.03	-0.076	-0.469		
12	6.46	20.00	0.32	34.06	0.219	-0.404		
13	7.00	60.00	0.12	58.99	0.000	-0.333		
14	9.58	40.00	0.24	50.42	0.130	-0.358		
15	10.29	40.00	0.26	51.05	0.149	-0.356		
16	9.96	20.00	0.50	81.97	0.405	-0.268		
17	11.88	40.00	0.30	60.27	0.191	-0.330		
18	11.88	40.00	0.30	60.27	0.191	-0.330		
19	11.88	40.00	0.30	60.27	0.191	-0.330		
20	10.41	20.00	0.52	51.39	0.428	-0.355		
21	7.54	20.00	0.38	55.27	0.276	-0.344		
22	11.37	20.00	0.57	52.44	0.479	-0.352		
23	5.92	40.00	0.15	34.56	0.033	-0.402		
24	8.61	60.00	0.14	43.78	0.029	-0.376		
25	9.06	20.00	0.45	50.04	0.357	-0.359		
26	3.48	60.00	0.06	17.41	-0.062	-0.451		
27	9.74	40.00	0.24	42.90	0.134	-0.379		
28	9.41	60.00	0.16	40.77	0.043	-0.385		
29	12.00	20.00	0.60	53.79	0.513	-0.348		
30	5.17	60.00	0.09	22.83	-0.032	-0.436		
31	22.95	40.00	0.57	88.08	0.485	-0.251		
32	11.88	40.00	0.30	60.27	0.191	-0.330		

It is worth noticing that the specific activity utility scores are all negative. This is because for all the architectures analyzed, the resulting specific activity falls below the minimum value reported by a potential "competitor".

4.4.3. COST MODEL

Rathi et al (2002) validated the attribute model obtained through RSM by producing lipase in a 14 liter fermenter with 10 liters of working volume using optimal conditions predicted by the model. The results and implications of this will be discussed in the following subsection. In this section, the result of the cost model developed assuming a 14L fermenter is presented.

So far, for the development of attribute and utility model, 5 design variables have been taken into account, namely: Glucose concentration, Palm oil concentration, incubation time, inoculum density and agitation. Incubation density was considered to have little effect on cost, thus neglected. For agitation and incubation time, similar to case study 1, power density was estimated. The table below shows the result of such calculation.

Table 21 – Case Study 2: Lipase power density and power consumption cost calculation

Average industrial electricity rate in Massachusetts is 12.57¢/kWh (Electricity Local, 2016).

RUN#	RPM	N	P	P/Vz	C/L	Incubation time	$/50M3
	1/min	1/seg	W	W/L		hours	
1	200	3.33	1.56	0.155	0.076	40.00	38.11
2	200	3.33	1.56	0.155	0.076	40.00	38.11
3	200	3.33	1.56	0.155	0.076	40.00	38.11
4	100	1.67	0.19	0.019	0.014	60.00	6.94
5	300	5.00	5.26	0.524	0.389	60.00	194.68
6	200	3.33	1.56	0.155	0.076	40.00	38.11
7	100	1.67	0.19	0.019	0.005	20.00	2.31
8	200	3.33	1.56	0.155	0.076	40.00	38.11
9	300	5.00	5.26	0.524	0.130	20.00	64.89
10	100	1.67	0.19	0.019	0.014	60.00	6.94
11	100	1.67	0.19	0.019	0.014	60.00	6.94
12	100	1.67	0.19	0.019	0.005	20.00	2.31
13	200	3.33	1.56	0.155	0.114	60.00	57.16
14	200	3.33	1.56	0.155	0.076	40.00	38.11
15	200	3.33	1.56	0.155	0.076	40.00	38.11
16	200	3.33	1.56	0.155	0.038	20.00	19.05
17	200	3.33	1.56	0.155	0.076	40.00	38.11
18	200	3.33	1.56	0.155	0.076	40.00	38.11
19	200	3.33	1.56	0.155	0.076	40.00	38.11
20	100	1.67	0.19	0.019	0.005	20.00	2.31
21	300	5.00	5.26	0.524	0.130	20.00	64.89
22	300	5.00	5.26	0.524	0.130	20.00	64.89
23	100	1.67	0.19	0.019	0.009	40.00	4.63
24	100	1.67	0.19	0.019	0.014	60.00	6.94
25	300	5.00	5.26	0.524	0.130	20.00	64.89
26	300	5.00	5.26	0.524	0.389	60.00	194.68
27	300	5.00	5.26	0.524	0.260	40.00	129.78
28	300	5.00	5.26	0.524	0.389	60.00	194.68
29	300	5.00	5.26	0.524	0.130	20.00	64.89
30	300	5.00	5.26	0.524	0.389	60.00	194.68
31	200	3.33	1.56	0.155	0.076	40.00	38.11
32	200	3.33	1.56	0.155	0.076	40.00	38.11

Glucose concentration and palm oil concentration cost for a $50m^3$ reaction was also estimated. Results are shown in the table below:

Table 22 – Case Study 2: Lipase architectures cost calculation

RUN#	Glu	Palm oil	Glucose			Palm oil			stirring	Total	
	mg/mL	% V/V	kg/L	kg/50M3	$/50M3	L/50M3	kg/50M3	$/50M3	$/50M3	$/50M3	glucose
1	8	2	0.008	400	140	1000	912.98	730.384	38.11	908.49	$0.35 kg
2	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	Palm oil
3	8	0	0.008	400	140	0	0	0	38.11	178.11	$0.8 kg
4	14	0	0.014	700	245	0	0	0	6.94	251.94	0.91298 kg/L
5	14	2	0.014	700	245	1000	912.98	730.384	194.68	1170.06	
6	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	
7	14	0	0.014	700	245	0	0	0	2.31	247.31	
8	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	
9	14	2	0.014	700	245	1000	912.98	730.384	64.89	1040.28	
10	14	2	0.014	700	245	1000	912.98	730.384	6.94	982.32	
11	2	0	0.002	100	35	0	0	0	6.94	41.94	
12	2	0	0.002	100	35	0	0	0	2.31	37.31	
13	8	1	0.008	400	140	500	456.49	365.192	57.16	562.35	
14	2	1	0.002	100	35	500	456.49	365.192	38.11	438.30	
15	14	1	0.014	700	245	500	456.49	365.192	38.11	648.30	
16	8	1	0.008	400	140	500	456.49	365.192	19.05	524.25	
17	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	
18	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	
19	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	
20	14	2	0.014	700	245	1000	912.98	730.384	2.31	977.70	
21	2	2	0.002	100	35	1000	912.98	730.384	64.89	830.28	
22	14	0	0.014	700	245	0	0	0	64.89	309.89	
23	8	1	0.008	400	140	500	456.49	365.192	4.63	509.82	
24	2	2	0.002	100	35	1000	912.98	730.384	6.94	772.32	
25	2	0	0.002	100	35	0	0	0	64.89	99.89	
26	14	0	0.014	700	245	0	0	0	194.68	439.68	
27	8	1	0.008	400	140	500	456.49	365.192	129.78	634.98	
28	2	2	0.002	100	35	1000	912.98	730.384	194.68	960.06	
29	2	2	0.002	100	35	1000	912.98	730.384	64.89	830.28	
30	2	0	0.002	100	35	0	0	0	194.68	229.68	
31	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	
32	8	1	0.008	400	140	500	456.49	365.192	38.11	543.30	

4.4.4. TRADESPACE

With the information obtained in the previous sections, a Multi-attribute score was calculated and plotted. For this, several scenarios were analyzed:

a) Both attributes 'productivity' and 'activity' having the same level of importance. Each attribute has a weight of 0.5.

b) Assuming that 'productivity' is expected to be further improved in bioreactor, 'activity' was given more importance. 'Productivity' weight was 0.2 and 'activity, weight was 0.8.

c) Assuming that specific activity of purified enzyme will be more than 1 times higher than in crude extract, 'productivity' was given more importance than 'activity'. 'Productivity weight was 0.8 and activity weight was 0.2.

The following figures shows the Tradespace for each of the scenarios. The tables with the calculation of each scenario are presented in appendix E.

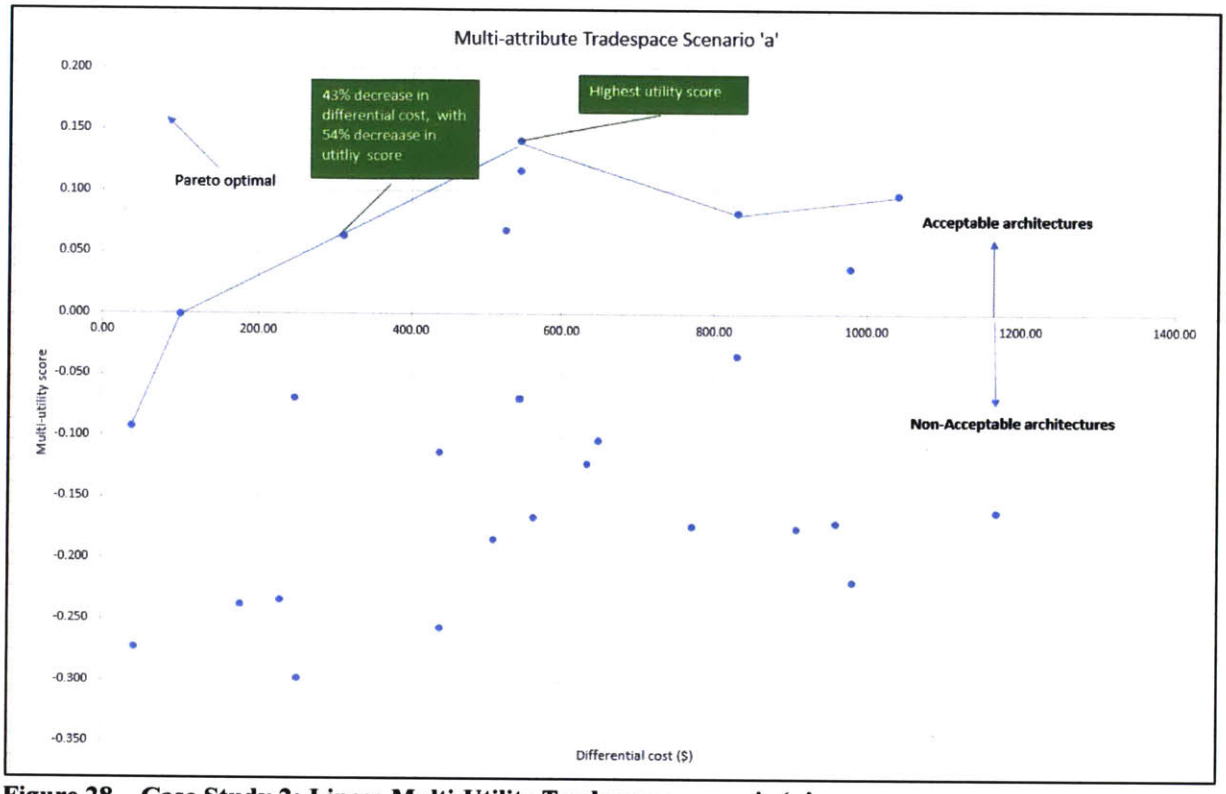


Figure 28 – Case Study 2: Lipase Multi-Utility Tradespace scenario 'a'

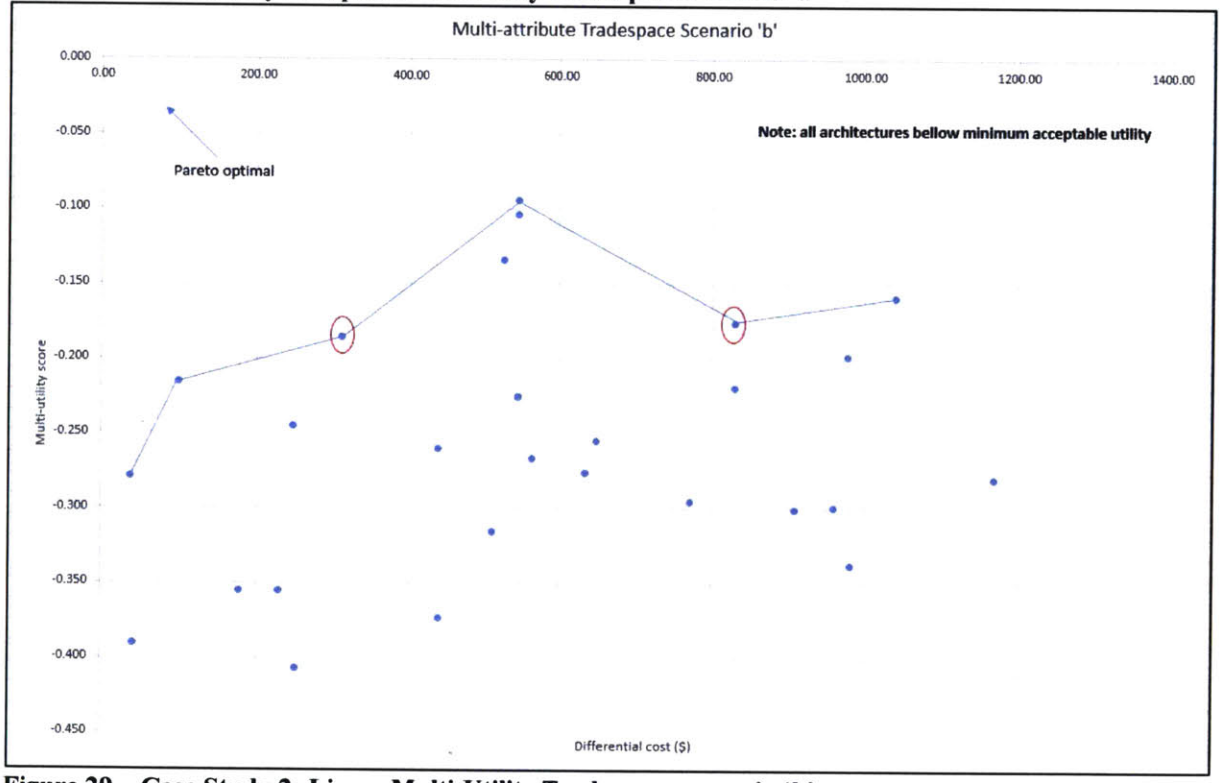


Figure 29 – Case Study 2: Lipase Multi-Utility Tradespace scenario 'b'

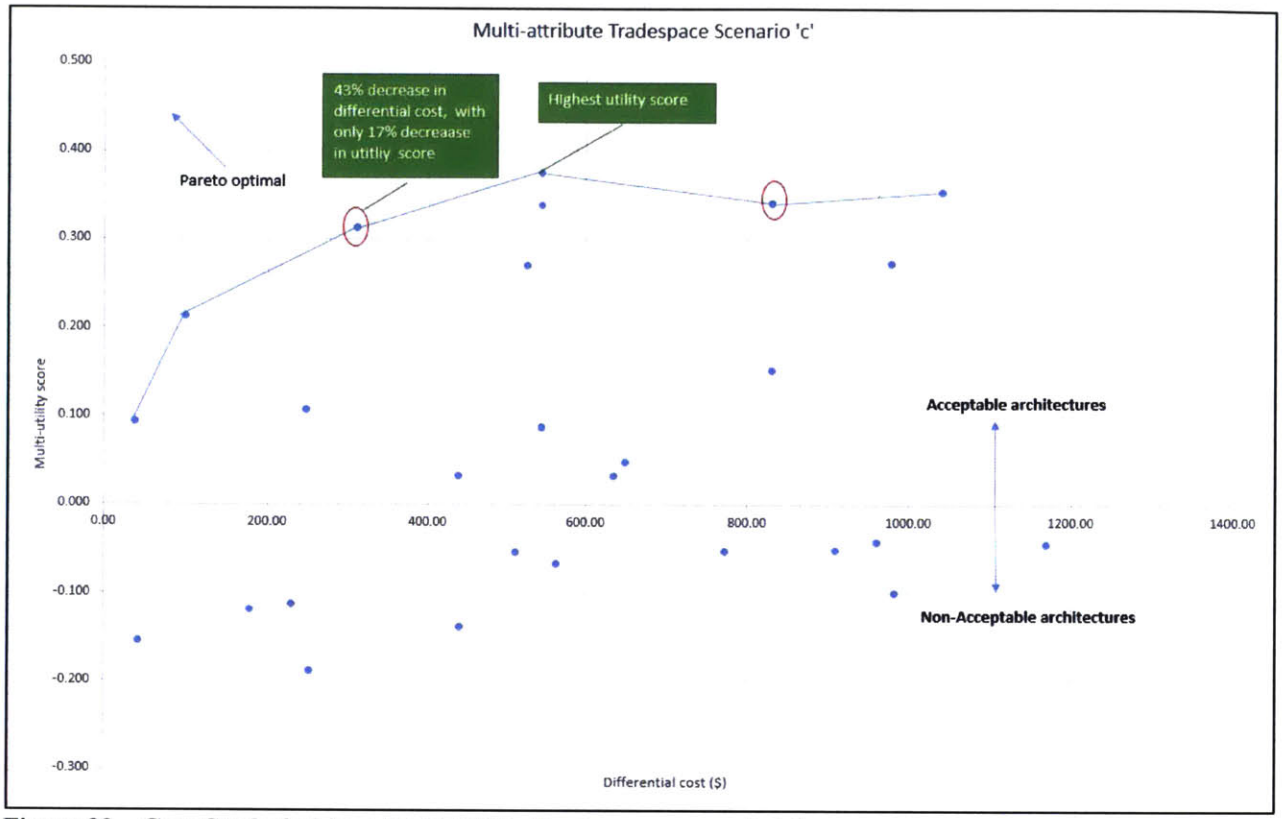


Figure 30 – Case Study 2: Lipae Multi-Utility Tradespace scenario 'c'

The first thing we can notice is that as the weight changes, the shape of the Tradespace changes slightly. However, the utility scores, and therefore the number of acceptable architectures, varies greatly. Such that in scenario 'b' ('Productivity' weight = 0.2, 'Activity' weight = 0.8) none of the architectures under analysis are acceptable. This is because utility is driven by the attribute 'activity' in this scenario and because lipase enzyme specific activity reported Rathi et al. (2002) are relatively low in comparison to Bhosale et al. (2016) and market standard. Therefore, based on this analysis Rathi et al (2002) would need to improve the specific activity of their enzyme by either increasing the amount of lipase produced per mg of total protein, developing a very efficient protein purification process, or both.

This case exemplifies the importance of correctly assigning weight to the different Single-Utility Functions (SUF). As observed, the different weights can lead to drastically different results. Scenario 'a' resulted in 7 acceptable architectures, scenario 'b' in none and finally scenario 'c' in 15.

When comparing non-dominated architectures; in other words, architectures on the Pareto frontier, two architectures stand out (shown in red circle in scenario 'b' and 'c'). As the weights of 'productivity' and 'activity' flip, these two architectures are responsible for the change in the shape of the Pareto frontier. The utility score of these two architectures in comparison to the rest of the architectures in the Pareto frontier drops significantly in scenario 'b', where the attribute 'activity' has more weight. This result suggests that these two architectures yields higher productivity (U/mL*hr) based on the increase of total protein and not lipase enzyme alone. If the increase of productivity observed was based on increase of lipase enzyme expression, then it is expected the utility score for specific activity (U/mg) would not drop, thus change of shape in the Pareto frontier would not occur. Therefore, this result highlights the importance of analyzing the complete bioprocess, as these two architectures might require a more efficient downstream processing.

Finally, similar to DHA case study, for scenario 'c' the architecture with the highest utility score might not necessarily be the optimal architecture, as the following architecture towards the left of the Pareto frontier costs 43% less, with only a 17% decrease in utility score. However, this analysis does not take into account the downstream processing. As discussed in the above paragraph, because the higher productivity of the architectures in red circle is apparently due to increase of overall protein production (not lipase alone), the downstream processing would have to compensate for this fact. Thus, if at the point of the analysis no downstream information is available, a "safer" approach would be scenario 'a' (a weight of 0.5 for each attribute). In this scenario, the conclusion derived in scenario 'c' is no longer true. In this case, the difference in cost between the highest utility architecture and the next architecture on the Pareto frontier towards the left (closer to Pareto optimal) remains 43%, but the decrease in utility score is 54%. Once again, this highlighting the importance of weight assignation and an analysis taking into consideration the complete process.

5. DISCUSSION AND CONCLUSION

It is to the author's understanding that the use of Tradespace, particularly MATE in the context of bioprocesses had not been proposed before, potentially due to the high complexity of biochemical processes, as it was demonstrated in the development of a DVM. Unlike other systems, in bioprocess systems it is non-trivial to link a given attribute to one (or a reduced number) of design variables. In order to overcome this issue, the present thesis proposed the use of DoE, specifically PB design, to identify most critical design variables followed by a RSM to develop single-utility curves. This new MATE-DoE methodology was tested in two case studies, proving the possibility of implementation.

These two case studies exemplified the benefit of an integrated development process, where several architectural designs are simultaneously assessed for more than one criteria. Unlike the traditional sequential process the "MATE-DoE" method proposed in this thesis avoids pre-mature focusing. It also allows a multi-objective, multi- attribute optimization. As a result, MATE-DoE method allows:

a) The exploration of several architectures (combination of process parameters) and comparison based on differential cost. Thus, allowing the identification of architectures with similar utility but different cost, giving rise to cost saving opportunities.
b) It allows a straightforward comparison to market standards and competitors. SUF (Single-Utility Functions) facilitate the comparison to exiting products, while at the same time they translate an attribute value into utility score. This last point allows the aggregation and simultaneous analysis of several attributes.
c) It assess a group of attributes, as opposed to an attribute in isolation. Customer decision to buy a given product is a complex process, where several attributes of the product are taken into account and trade-offs are made. The proposed MATE-DoE method takes this complexity into account. The weighted sum method used to develop a MUF (Multi-Utility Function) allows decision maker to score architectures based on how well they perform in achieving a group of attributes according to customer preference (in terms of which attributes and in order of importance). It allows the aggregation of several attributes into a dimensionless parameter, referred to as utility.

It is important to point out that the aim of the MATE-DoE method is to complement existing process simulation software. It combines and links bioprocess optimization practice with market

research results. Thus, it promotes multi-disciplinary, cross-functional team work. Furthermore, the main purpose of MATE-DoE, would be to analyze different scenarios at a time, comparing its potential total, or differential cost. As shown in both case studies, this would potentially support the identification of architectures with slightly lower utility but allowing considerable savings in cost. Furthermore, if a specific market segment and price can be identified, a Tradespace of profitability vs. utility would clearly show the tradeoffs between this two objectives.

While the proposed method was successfully used in two case studies, in the implementation of MATE-DoE in both case studies some challenges where encountered:

1. **Incomplete experimental data.** Due to time constrains, experimental data was obtained from published papers. Thus, the analysis had to be adapted to the published information as opposed to carrying out experiments to solve the question of what are the most important design variables for the attributes identified. Therefore, experimental results measuring oxidation levels of DHA was not found. Instead, in both cases the attribute 'productivity' was analyzed. This is because the objective of process development is typically to increase productivity. Furthermore, for case study 2, ideally, the experiments would have been carried out in small fermenters, as opposed to flask experiments, as it is well known that due to superior aeration conditions, productivity in bioreactors is higher.
2. **Assumption in minimum, maximum utility value and utility weight.** It is important to point out that in a real-life scenario, surveys, interviews and deeper market analysis need to be performed to identify reasonable maximum, minimum, utility values and weights. This process might be long and would potentially require several discussion with the development team, as a 'right' values does not exist; rather it is an agreement and an educated guess based on market data. For both case studies, an effort was made to find commercial DHA and lipase data. However, interviews with experts and continuous discussion with the development team were beyond the scope of the present thesis.
3. **Research stage of experimental data.** Published data in scientific journals are typically in research stage, often times conducted at a university laboratory. To be more accurate, and to be able to compare to maximum and minimum attribute values, the experimental data should be beyond the research stage and rather in process development stage. This might explain, for example, the low 'specific' activity' values reported by Rathi et al (2002) for lipase enzyme.

The above challenges highlight the importance of three elements for the successful implementation of the proposed "MATE-DoE" method:

a) **Thorough market analysis to identify customer attribute preference.** As proven in the second case study, the weighted sum method used to derive a MUF (Multi-Utility Function) score can yield different results when weights are assigned differently. For example Multi-utility Tradespace scenario 2 of lipase enzyme case study, resulted in zero acceptable architectures. This subjective scoring system can hugely misguide the analysis if not performed appropriately.

b) **Thorough market analysis to identify attribute maximum and minimums.** In order for the MATE-DoE Tradespace analysis to be significant, the appropriate identification of attribute maximum and minimum acceptable values for each Single-Utility Function is essential. Otherwise, what appears to be a high utility score architecture can in reality be a low performing architecture.

c) **Coordination, analysis and discussion with the process development team.** The identification of appropriate attribute maximum, minimum and weight, involves not only the marketing team, but also the process development and management team. As mentioned before, a "right/correct" value does not exist. But rather it is an agreement based on market and existing process information.

   Also, due to the number of experiments needed to be carried out in DoE (even though less than one-at-a-time experiments, nonetheless still a considerable number of experiments is required), a good coordination with the process development team is essential to avoid unnecessary experiment repetition. Ideally, a good planning would result in a coordinated effort to collect information for several attributes at once, and not repeat experiments (i.e.: one set of experiments to measure productivity, oxidation level, purity, etc.)

Finally, due to the great number of experiments and market analysis involved, further studies to simplify the method are recommended. Potential ways of simplifying the method are:

1) **MATE-DoE applied only on the fermentation and downstream steps.** While upstream processes are very important in the process development, they have little impact on cost. MATE-DoE Tradespace is meant to help analyze the tradeoffs between attributes (productivity, activity, oxidation level, purity, etc) and cost (total cost, differential cost,

profitability, etc.). Thus, there seems to be no point on carrying out MATE-DoE Tradespace in upstream process.

2) **PB design as a go/ not go decision for MATE-DoE.** After identifying critical design parameters for a given attribute, a cost calculation assuming all the identifying parameters in their high (+1) and low value (-1) can be carried out. The estimated cost in a scale up scenario can be compared. If the difference (at scale) between an architecture with all the design variables in +1 and an architecture with all the design variables in -1 is not significant, then a MATE-DoE Tradespace analysis would not be recommended.

3) **Coding to create more architectures.** Once single-utility functions are developed, more architectures (within the +1 and -1 values of each design variable) can be predicted using software such as MATLAB ®. In this way more architectures can be identified along the Pareto frontier without carrying out the required experiments. Also, this would avoid "artificial clustering" due to DoE. For example, in case study 1 (DHA), the trade space seems to suggest 5 clusters as shown in the figure bellow:

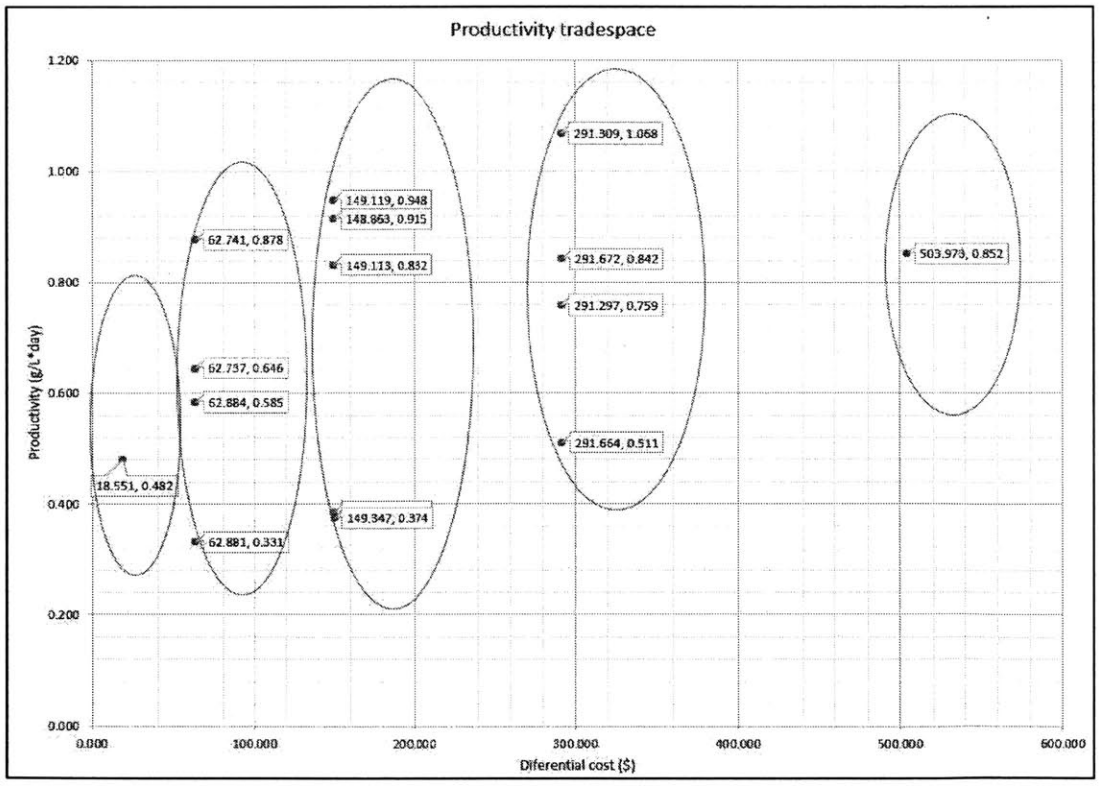


Figure 31 – "Artificial clustering" in Tradespace.

What appears as 5 distinctive clusters is in reality the 5 levels analyzed in the Central Composite Circumscribed DoE. Because in case study 1, a simplified analysis taking into consideration only one utility (productivity in this case) and one cost driver (i.e. power density) the architectures are clustered and correspond to the 5 levels of the DoE design. This phenomenon was not observed in case study 2, as more than one utility and cost driver was taken into account in the analysis. Computer simulations can further prevent this artifact. By using the corresponding Single and Multiple Utility Functions, computer simulations can populate the Tradespace, showing a greater number of possible architectures, avoiding this "artificial clustering" created by experimental results.

Finally, further studies incorporating downstream process to further assess the applicability of the proposed method are required to better understand its applicability. As it was shown in case study 2, the analysis of the complete bioprocess is essential for an accurate assessment.

In conclusion, the proposed "MATE-DoE" could potentially contribute and complement existing bioprocess simulator software by performing a more holistic, integrated analysis and promote early discussion between marketing and process development team. However, this method should be used with caution, as inaccurate input data could yield misguiding results. Particularly, the weighted sum method to produce Multi-Utility Functions (MUFs) has to be appropriately performed. Further studies to simplify the proposed method are recommended. Also, case studies analyzing the complete bioprocess are required to further assess the applicability of the "MATE-DoE" method proposed.

6. LIST OF REFERENCES

Amore, A., & Faraco, V. (2016). Enzymes for Food and Beverage Industries: Current Situation, Challenges and Perspective. In R. Rai, *Advances in Food Biotechnology* (pp. 165-183). Chinchester: Wiley & Sons.

Andualema, B., & Gessesse, A. (2012). Microbial Lipase and their Industrial Applications: Review. *Biotechnology*, 100-118.

Aravindan, R., Anbumathi, P., & Viruthagiri, T. (2007). Lipase applications in food industry. *Indian Journal of Biotechnology*, 141-158.

Bernstain, J. (1998). *Design Methods in the Aerospace Industry: Looking for Evidence of Set-Based Practices.* Cambridge: Master of Science Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology.

Bhosale, H., Shaheen, U., & Kadam, T. (2016). Characterization of a Hyperthermostable Alkaline Lipase from Bacillus sonorensis 4R. *Enzyme Research*, 1-12.

Box, G., Hunter, J., & Hunter, W. (2005). *Statistics of Experimenters: Design, Innovation and Discovery* . Hoboken, NJ: John Wiley & Sons, Inc.

Chisti, Y. (2010). Fermentation Technology. In W. Soetaert, & E. J. Vandamme, *Industrial Biotechnology. Sustainable Growth and Economic Success* (pp. 149-171). Verlag: Wiley-VCH.

Crawley, E., Cameron, B., & Selva, D. (2016). *System Architecture.* Hoboken, NJ: Pearson Higher Education, Inc. .

Darvasula, R. V., Darvasula V., S. R., & Rao, V. (2013). Microalgal Biotechnology: Today's (Green) Gold Rush. In *Biotechnology Application of Microalgae Biodiesel and Value-Added Products* (pp. 201-228). Boca Raton: CRC Press.

De Swaaf, M., De Rijk, T., Eggink, G., & Sijtsma, L. (1999). Optimisation of docosahexaenoi acid production in batch cultivation by Crypthecodinium cohnii . *Journal of Biotechnology* , 185-192.

Demain, A. L. (2010). History of Industrial Biotechnology. In W. Soetaert, & E. L. Vandamme, *Industrial Biotechnology. Sustainable Growth and Economic Success.* (pp. 17-68). Weinheim: Wiley.

Dori, D. (1995). Object-Process Analysis:Maintaining the Balance between System Strucure and Behaviour". *Journal of Logic and Computation*, 227-249.

Dori, D. (2002). *Object-Process Methodology - A Holistic System Paradigm.* Berlin Heidelberg: Springer .

Electricity Local. (2016). *Electricity Local.* Retrieved from Electricity Local: http://www.electricitylocal.com/states/massachusetts/

Erickson, B., Nelson, J. E., & Winters, P. (2012). Perspectiveon opportunities in industrial biotechnology in renewable chemicals. *Biotechnology Journal*, 176-185.

EuropaBio and ESAB. (2013, July). *Today's applicaitons: Food and Feed.* Retrieved from Bio-economy: http://www.bio-economy.net/applications/applications_food_and_feed.html

Ferreira-Dias, S., Sandoval, G., Plou, F., & Valero, F. (2013). The potential use of lipases in the production of fatty acid derivatives for the food and nutraceutical industries. *Electronic Journal of Biotechnology.*

Freitas, A. C., Rodrigues, D., Rocha-Santos, T. A., Gomez, A. M., & Duarte, A. C. (2012). Marine Biotechnolgy Advances Towards Applications in New Funcitonal Food . *Biotechnology Advances*, 1506-15.

Frost and Sullivan. (2014, July 14). Global Omega-3 and Omega-6 Polyunsaturated Fatty Acid Ingredients Market (NE34-01). Frost and Sullivan.

Gill, N., Appleton, M., Baganz, F., & Lye, G. (2008). Quantification of Power Consumption and Oxygen Transfer Characteristics of a Stirred Miniature Bioreactor for Predictive Fermentation Scale-Up. *Biotechnology and Bioengineering*, 1-12.

Guesnet, P., & Alessandri, J.-M. (2011). Docosahexaenoic acid (DHA) and the developing central nervous system (CNS)- Implications for dietary recommendations. *Biochimie*, 7-12.

Hallmann, A. (2007). Algal Transgenics and Biotechnology. *Transgenic Plant Journal*, 81-98.

Harun, R., Singh, M., Forde, G., & Danquah, M. (2010). Bioprocess Engineering of Microalgae to Produce a Variety of Consumer Products . *Renowable Sustainable Energy Rev*, 1037-47.

Heerlen. (2010, December 21). *www.dsm.com*. Retrieved from DSM: http://www.dsm.com/corporate/media/informationcenter-news/2010/12/57-10-dsm-to-acquire-martek-to-add-new-nutrition-growth-platform.html

Heinzle, E., & Hungerbuhler, K. (1997). Integrated Process Development: The Key to Future Production of Chemicals. *Chimia*, 176-183.

Heinzle, E., Biwer, A. P., & Cooney, C. (2006). *Development of Sustainable Bioprocesses-Modeling and Assessment.* Chinchester: John Wiley & Sons.

Holland, F., & Chapman, F. (1966). *Liquid Mixing and Processing in Stirred Tanks.* New York: Reinhold Publishing Corp.

Iacurci, J. (2014, August 14). *Nature World News.* Retrieved from www.natureworldnews.com: http://www.natureworldnews.com/articles/8554/20140814/could-omega-3-craze-destroy-fish-populations.htm

iGEM2010. (n.d.). *iGEM UCL*. Retrieved from iGEM UCL: http://2010.igem.org/Team:UCL_London/Fermenter_Mechanics

Jaeger, K., & Eggert, T. (2002). Lipaes for biotechnology . *Current Opinion in Biotechnology*, 390-397.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York: Wiley.

Keeney, R., & Raiffa, H. (1993). *Decisions with Multiple Objective-REferences and Value Tradeoffs.* Cambridge, England, UK: Cambridge University Press.

Khosravi-Darani, K., Koohy-Kamaly, P., Nikoopour, H., & Zeinab Asadi, S. (2016). Production of Single-Cell Oil Containing Omega-3 ad Omega-6 Fatty Acid. In R. Rai V, *Advances in Food Biotechnology* (pp. 369-380). Chichester: Wiley & Sons.

Lebed, P. J., Potvin, S., Lariviere, D., & Dai, X. (2014). Optimization of solid phase extraction chromatogrphy for the separation of Np from U and Pu using experimental desing tool in complex matrices. *Analytical Methods*, 139-146.

Mackertich, N., & Kraus, P. (2008, October 15). Exploring and Evaluating the Product Performance- Cost Design Trade Space. *Global Lean, Six Sigma & Business Improvement Summit.* Raytheon Company.

Mandenius, C.-F., & Brundin, A. (2008). Review: Biocatalysis and Bioreactor Design . *Biotechnology Progress*, 1191-1203.

Markets and Markets. (2015). *Lipase Market by Source (Microbial Lipases, Animal Lipases), Application (Animal Feed, Dairy, Bakery, Confectionery, Others), & by Geography*

(North America, Europe, Asia-Pacific, Latin America, RoW) - Global Forecast to 2020. Magarpatta: Markets and Markets.

McGovern, P., Zhang, J., & J.G, T. (2004). Fermented beverages of pre- and proto-historic China. *Proceedings of the National Academy of Science*, 17593-17598.

Miller, J. (2013, March 11). *Experimental design and optimization.* Retrieved from www.rsc.org: http://www.rsc.org/images/Experimental-design-and-optimisation-4-Plackett-Burman-designs-55_tcm18-232212.pdf

Montgomery, D. (1997). *Design and Analysis of Experiments.* New York: John Wiley & Sons, Inc.

Mordor Intelligence. (2015). *Global Food Enzymes Market-Groth, Trends and Forecast (2015-2020).* Lyon: Reportlinker .

Nikou, T., & Klotz, L. (2014). Application of multi-attribute utility theory for sustainable energy decisions in commercial building. *Smart and Sustainable Built Enviroment* , 207-222.

NIST/SEMATECH. (2013, 10 30). *e-Handbook of Statistical Methods.* Retrieved from Engineering Statistics Handbook: http://www.itl.nist.gov/div898/handbook/pri/section3/pri3361.htm

Ogle, R. A., Dee, S. J., & Cox, B. L. (2015). Resolving inherently safer desing conflicts with decision analysis and multi-attribute utility theory . *Process Safety and Enviromental Protection* , 61-69.

Owen, M., Luscombe, C., Lai, L.-W., Godbert, S., Crookes, D., & Emiabata-Smith, D. (2001). Efficiency by design: optimisation in process research. *Organic Process Research and Development*, 308-323.

Petrides, D. (2000). *Bioprocess Design.* Scotch Plains: Intelligen Inc. .

Petrides, D., Carmichael, D., Siletti, C., & Koulouris, A. (2014). /biopharmaceutical Process Optimiztion with Simulation and Scheduling Tools. *Bioengineering* , 154-187.

Plackett, R., & Burman, J. (1946). The Design of Optimum Multifactorial Experiments. *Biometrika* , 305-325.

Plasun, R. (n.d.). *Institute of Microelectronics.* Retrieved from http://www.iue.tuwien.ac.at/: http://www.iue.tuwien.ac.at/phd/plasun/node32.html

Pulz, O., & Gross, W. (2004). Valuable producst from biotechnology of microalgai. *Appl Microbiol Biotechnology*, 635-648.

Rahman, M. (2016). Application of Metabolic Engineering in Industrial Fermentative Process. In R. Rai, *Advances in Food Biotechnology* (pp. 221-242). Chichester: John Wiley & Sons.

Raiffa, H. (2002). Decision analysis: a personal account ofhow it got started and evolved. *Operations Research* , 179-185.

Rathi, P., Goswami, V., Sahai, V., & Gupta, R. (2002). Statistical medium optimization and production of hyperthermostable lipase from Bukholderia cepacia in bioreactor. *Journal of Applied Microbiology*, 930-936.

Rathi, P., Saxena, R., & Gupta, R. (2001). A novel alkaline lipase from Burkholderia cepia for detergent formulation. *Process Biochemistry*, 187-192.

Ross, A. M. (2003). *Multi-Attribute Tradespace Exploration with Concurrent Desing as a Value-centric Framework for Space System Architectue and Design* . Cambridge, MA: Aeronautics and Astronautics Technology and Policy Program, Massachusetts Institute of Technology. Dual Master of Science Thesis.

Ross, A. M., & Hasting, D. E. (2005). *The Tradepace Exploration Paradigm.* Rochester, NY: INCOSE International Symposium.

Ross, A., & Hasting, D. (2002). Multi-Attribute TRadespace Exploration in Space System Design. *53rd International Astronautical Congress- The World Space Congress.* . Houston, TX: IAF IAC-02-U3.03.

Shanklin, T., Roper, K., Yegneswaran, P., & Marten, M. R. (2001). Selection of Bioprocess Simulation Software for Industrial Applications. *Biotechnology and Bioengineering*, 483-489.

Sharma, R., Christi, Y., & Banerjee, U. C. (2001). Production, purification, characterization and application of lipases. *Biotechnology advances*, 627-662.

Simpson, E. (2005, 11 15). *Bio- Biotechnology Innovation Organization.* Retrieved from What is Industrial Biotechnology?: https://www.bio.org/articles/what-industrial-biotechnology

Song, X., Zhang, X., Guo, N., Zhu, L., & Kuang, C. (2007). Assessment of marine thraustochytrid Schizochytrium limacinum OUC88 for mariculture by enriched feeds. *Fisheries Science*, 565-573.

Song, X., Zhang, X., Kuang, C., Zhu, L., & Guo, N. (2007). Optimization of fermentation parameters for the biomass and DHA production of Schizotrium limacicum OUC88 usign resposnse surface methodology. *Process Biochemistry*, 1391-1397.

Spaulding, T. J. (2003, May 9). Tools for Evolutionary Acquisition: A Stuydy of Multi-Attribute Tradespace Exploration (MATE) Applied to the Space Based Rada (SBR). Cambridge, Massachusetts, United States: Aeronautics and Astronautics SM, Massachusetts Institute of Technology.

Spaulding, T. J. (2003, May 9). Tools for Evolutionary Acquisition: A Stuydy of Multi-Attribute Tradespace Exploration (MATE) Applied to the Space Based Rada (SBR). Cambridge, Massachusetts, United States: Massachusetts Institute of Technology.

Spolaore, P., Joannis-Cassan, C., Duran, E., & Isambert, A. (2006). Commercial applications of microalgae. *Journal of Bioscience and Bioengineering*, 87-96.

Stowe, R., & Mayer, R. (1966). Efficient screening of process variables. *Ind Eng Chem*, 36-40.

The Economist . (2009, June 4). Third time lucky. *The Economist* .

Triantaphyllou, E., Shu, B., Nieto Sanchez, S., & Ray, T. (1998). Multi-Criteria Decision Making: An Operations Research Approach. In J. G. Webster, *Encyclopedia of Electrical and Electronics Engineering* (pp. 175-186). New York: John Wiley & Sons.

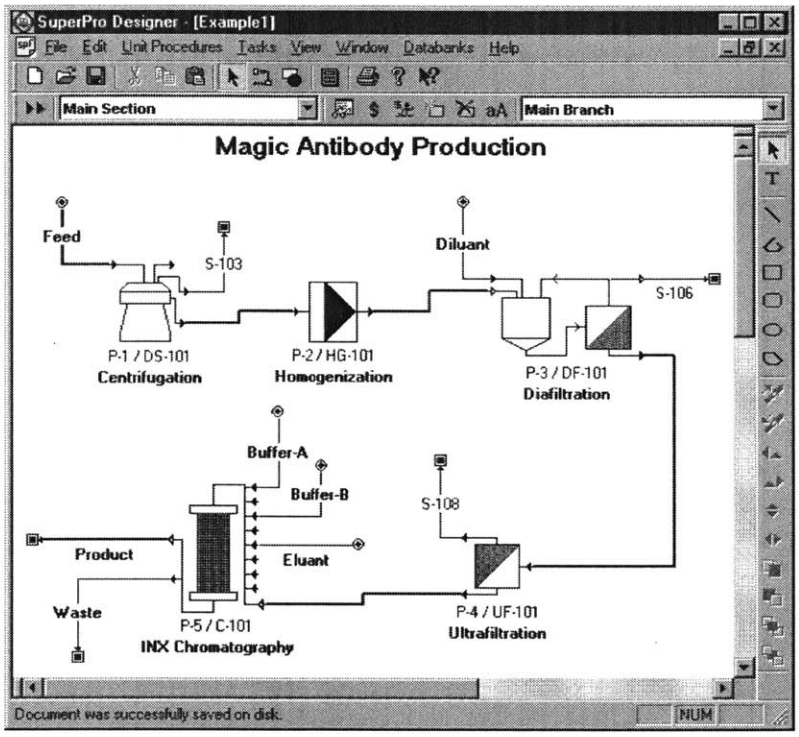
Vermelho, A. B., Cardoso, V., Pires Nascimento, R., Pinheiro, A. S., & Rodriguez, A. (2016). Application of Microbial Enzymes in the Food Industry. In R. Rai, *Advances in Food Biotechnology* (pp. 105-125). Chinchester: Wiley & Sons.

Vigani, M., Parisi, C., Rodriguez-Cerezo, E., Barbosa, M. J., Sijtsma, L., Ploeg, M., & Enzing, C. (2015). Food and fedd products fom micro-algae: Market opportunities and challenges for the EU. *Trends in Food Science and Technology* , 81-92.

Von Neumann, J., & Morgenstern, O. (1944/1947/1953). *Theory of Games and Economic Behaviour.* Princeton, NJ: Princeton University Press.

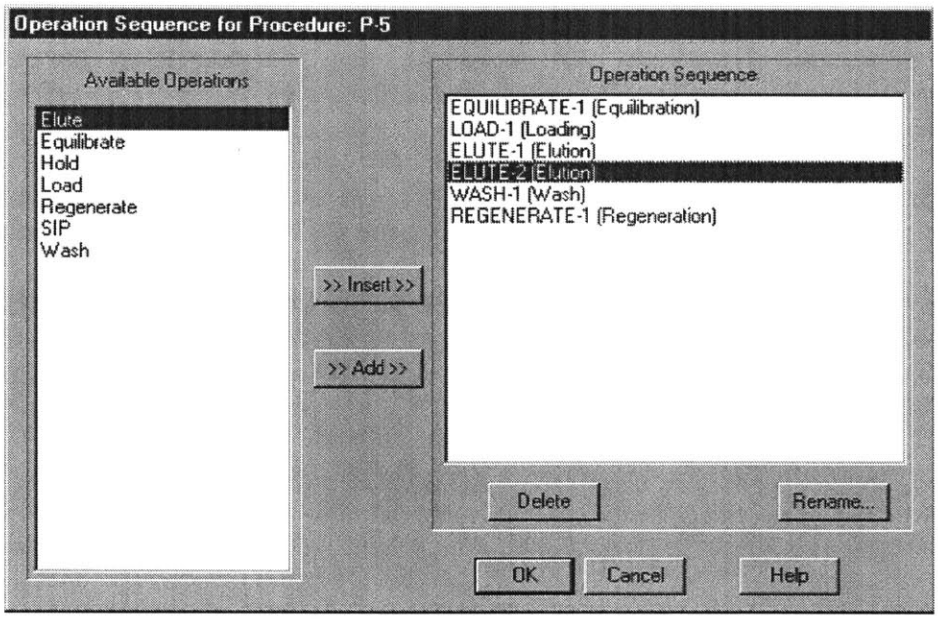
APPENDIX A – SUPERPRO SOFTWARE OVERVIEW

To model an integrated process on the computer using a simulator, a flowsheet that represents the overall process is first developed. The figure bellow for instance displays the floowsheet of a hypothetical process in the main window of SuperPro Designer. The flowsheet is developed by putting together the required unit operations and joining them with material flow streams.

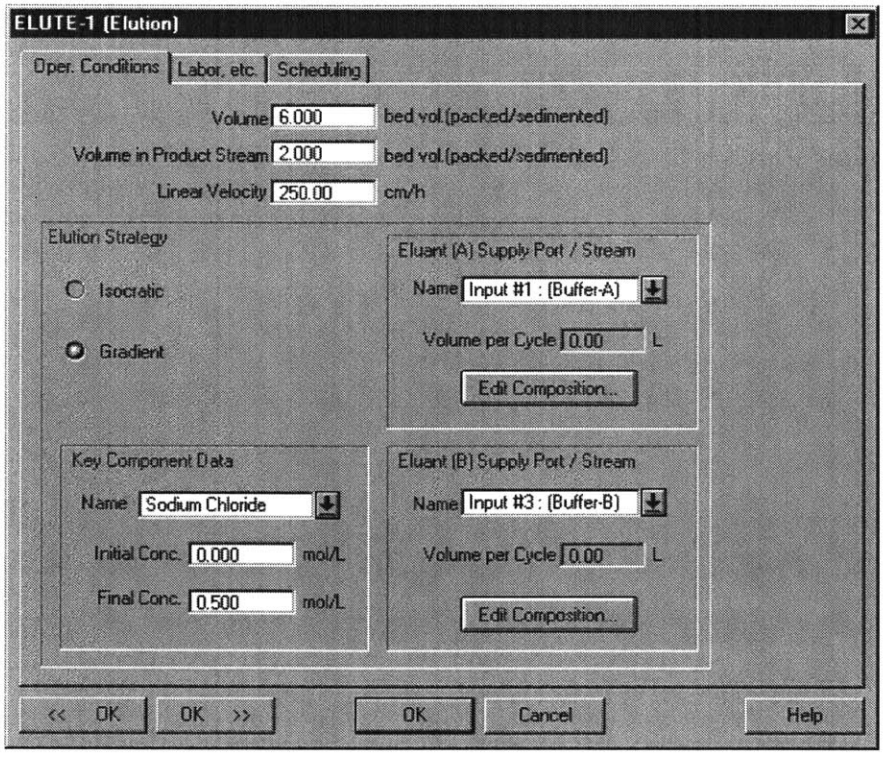


Next, the user initialize the flowsheet by registering (selecting from the component database) the various materials that are used in the process and specifying operating conditions and performance parameters for various operations.

In continuous operations, a piece of equipment performs the same action all the time. In batch processing, on the other hand, a piece of equipment goes through a cycle of operations. For instance, a typical chromatography cycle includes *equilibration, loading, washing, elution and regeneration*. In SuperPro Designer, the set of operations that compromise a processing step is called "unit procedure" (as opposed to "unit operation"). Each unit procedure contains individual tasks (e.g., equilibration, loading, etc,.) called operations.



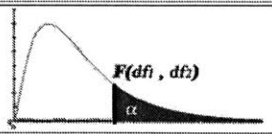
For every operation within a unit procedure, SuperPro includes a mathematical model that performs material and energy balance calculations. Before any simulation can be done, the user must initialize the various operations by specifying operating conditions and performance parameters, as shown below:



Source: (Petrides, Bioprocess Design, 2000)

APPENDIX B – F- DISTRIBUTION

F Table for $\alpha = 0.05$



/	df_1=1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
df_2=1	161.4476	199.5000	215.7073	224.5832	230.1619	233.9860	236.7684	238.8827	240.5433	241.8817	243.9060	245.9499	248.0131	249.0518	250.0951	251.1432	252.1957	253.2529	254.3144
2	18.5128	19.0000	19.1643	19.2468	19.2964	19.3295	19.3532	19.3710	19.3848	19.3959	19.4125	19.4291	19.4458	19.4541	19.4624	19.4707	19.4791	19.4874	19.4957
3	10.1280	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432

Source: http://www.socr.ucla.edu/applets.dir/f_table.html#FTable0.05

APPENDIX C – STAKEHOLDERS NEED DESCRIPTION

The following table lists and describes the stakeholders taken into consideration in this analysis:

Stakeholder	Description
Project/biotech startup	Biotech startup developing and owners of the bioprocess.
Community	Community provides qualified workforce for the startup
Partner 1	Bigger ingredient company or a CMO (Contract Manufacturing organization) with industrial capabilities to produce the food ingredient at scale. For example: a given new product development department within a bigger company, such as DSM, BASF, Novozyme, or an external CMO.
Partner 2	Bigger ingredient company in charge of commercializing the ingredient. If it is the same company as Partner 1 and there is a licensing agreement in place, then the difference between sales and finance department will be made, naming them Partner 2.a and Partner 2.b respectively.
Food manufacturing company	Food manufacturing company that uses the food ingredient to manufacture their products. For example: Nestle, P&G, others.
Distributors, Wholesaler and Retailer	In this analysis these three stakeholders have been aggregated, since they have similar needs. As a group, this distribution channel has the objective of selling food products to the consumers. Example of retailers: Costco, Target, Walmart, etc.
Consumer	The end user, who buys and consumes the food product.
Private investors	This could be venture capital, or angel investors of the biotech startup. But it could also be a private company investing in a specific project. However, independent of the type of investor, this is an entity that provides cash so that the project can be carried out.
Government investment	It can be investment in the form of grants, or other incentives such as tax incentives or loans. Examples are federal agencies such as NIH (National Institute of Health), NSF (The National Science Foundation) and programs such as Small Business Innovation Research (SBIR) and the Small Business Technology Transfer (STTR).
Regulators	Governmental entities in charge of enforcing the law and ensuring the safety of the ingredient and food for the population and the environment. For example: FDA, EPA
NGOs	Non-Governmental entity that advocate for food safety and communicate their view on healthy food and food ingredients. They create change by raising awareness about certain topics. Some controversial topics NGOs might advocate for within food industry might be: 1) traceability 2) labelling 3) health claims 4) environmental friendly and sustainabilityprocess. However, one of the most important topic in this case might be transgenic ingredients. Some examples of NGOs in the United States are: Center for food safety (http://www.centerforfoodsafety.org/), Food and Water watch (http://www.foodandwaterwatch.org/), Food Mythbusters (http://foodmyths.org/), etc.
Healthcare community	Similar to NGOs they also advocate for food and food ingredient safety. However, in this case rather than an established organization we are referring to the physicians, nurses and other healthcare personnel whose opinion on health safety of a given food and/or food ingredient might pursue the customer to consume or not a given product.

Raw material, utility and consumable suppliers	Here the suppliers were also aggregated, as their needs (mostly revenue) have little influence in the analysis of food ingredient specification. Instead quality specifications are impose on these suppliers to meet food ingredient quality needs. Also, difference between the types of suppliers was omitted for the same reason.

It is worth noticing that stakeholders that has little influence in the 'project/biotech start up' were not taken into account. Thus, farmers and other raw material providers for the food manufacturer are beyond the scope of this analysis. Also, some stakeholders such as 1) other ingredient companies that reflects the market share, or have a positive role in growing the market; 2) packing companies, important for the food manufacturing company, 3) academia, from which technical knowledge and new ideas for ingredient bioprocess production may emerge, among others stakeholder are not taken into account, as they have little influence in determining product specification, which is the objective of the present stakeholder analysis.

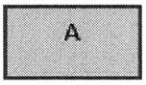
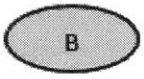
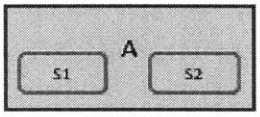
It is also important to highlight that in this analysis only the food and beverage market segment is taken into account. In other segments such as supplement or animal feed, the stakeholders and the nature of their interaction might be different. For instance, for feed industry, 'pellet manufacturer' and 'farmer' would need to be considered and the distribution channels would differs greatly.

Finally, the nature of the interaction might be different and case specific, depending on the kind of agreement the biotech start-up is able to negotiate. For instance, for lipase enzyme, the biotech company could choose to license the process to a big company such as Novozyme (who then becomes 'Partner 1' and 'Partner 2') or raise funds to produce it themselves and sell it to the ingredient to a bigger company (in this case 'Partner 1' would be a CMO and 'Partner2' the bigger company buying the ingredient). Since the specific type of agreement cannot be defined beforehand and since the objective is to identify product specifications, it is expected that the characteristics of the agreement will not impact the underlying need of each stakeholder. Thus, it was assumed for this stakeholder analysis that the biotech startup had a licensing and royalty agreement with a bigger firm, even though the author recognizes that this is not always the case.

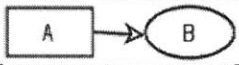
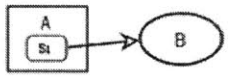
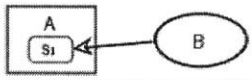
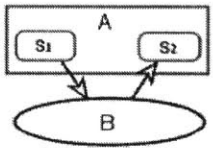
The following step is to describe each stakeholder's needs.

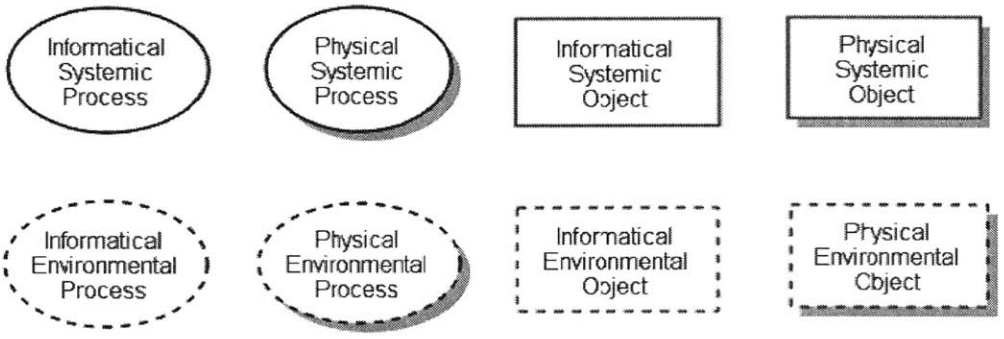
Stakeholder	Need	From
Project/biotech startup	Workforce	Community
	Licensing fee	Partner 2.b
	Royalty on sales	Partner 2.b (however, only possible if sales is received form Food manufacturing company)
	Investment	Private and government investment
	Equipment and raw material	Raw material suppliers
	Regulatory approval	Granted by regulators
Partner 1 (process development department)	Well defined process parameters	Biotech startup
	Salary	Partner 2.b
Partner 2.a (sales department)	Food ingredients	Partner 1
	Payment	Food manufacturing company
	Salary	Partner 2.b
Partner 2.b (finance department)	Income	Partner 2.a
Food manufacturing company	Food ingredients	Partner 2.a
Distributors, Wholesaler and Retailer	Food products	Food manufacturing company
	Payment	Consumer
Consumer	Food product	Retailers
	Recommendation	Healthcare community and NGOs
Private investors	ROI	Biotech Statup.
Government investment	Project report	Biotech Startup
Regulators	Safe and environmentally friendly process	Biotech Startup, Partner 1 and Food manufacturer
NGOs/local communities	Safe and environmentally friendly product and process	Biotech Startup, Partner 1 and Food manufacturer
Healthcare community	Safe and Healthy	Biotech Startup, Partner 1 and Food manufacturer
Raw material and consumable suppliers	Revenue	Biotech Startup

APPENDIX D – OBJECT PROCESS LANGUAGE

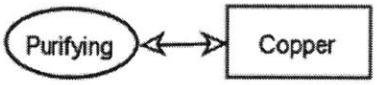
Symbol		Name: Definition	OPL	Comments
Things		Object **A:** A thing that exists	**A** is physical (and environmental).	**A** is informatical and systemic by default.
		Process **B:** A thing that transforms (generates, consumes, or changes the state of an object).	**B** is physical (and environmental).	**B** is informatical and systemic by default.
		State: A situation of an object.	**A** is s1. **A** can be s1 or s2. **A** can be s1, s2, or s3.	Always within an object.

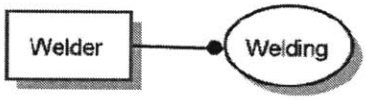
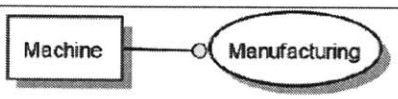
Symbol	Name	OPL	Allowed Source-to-Destination connections	Semantics/ Effect on the system flow/ Comments
▲	Aggregation-Participation	**A** consist of **B**.	Object-Object Process- Process	Whole -Part
	Exhibition-Characterization	**A** exhibits **B**.	Object-Object Object-Process Process-Object Process- Process	
△	Generalization-Specialization	**B** is an **A**. (objects) **B** is **A**. (processes)	Object-Object Process- Process	
	Classification-Instantiation	**B** is an instance of **A**.	Object-Object Process- Process	
	Tagged structural links: Unidirectional Bidirectional	According to text added by user	Object-Object Process- Process	Describes structural information.

Name	Symbol	OPL	Semantics
Consumption Link		**B** consumes **A**.	Process B consumes Object A.
State-Specified Consumption Link		**B** consumes **s1 A**.	Process B consumes Object A when it is at State s1.
Result Link		**B** yields **A**.	Process B creates Object A.
State-Specified Result Link		**B** yields **s1 A**.	Process B creates Object A at State s1.
Input-Output Link Pair		**B** changes **A** from **s1** to **s2**.	Process B changes the state of Object A from State s1 to State s2.
Effect Link		**B** affects **A**.	Process B changes the state of Object A;



Informatical Systemic Process is an informatical and systemic process.
Physical Systemic Process is a physical and systemic process.
Informatical Systemic Object is an informatical and systemic object.
Physical Systemic Object is a physical and systemic object.
Informatical Environmental Process is an informatical and environmental process.
Physical Environmental Process is a physical and environmental process.
Informatical Environmental Object is an informatical and environmental object.
Physical Environmental Object is a physical and environmental object.

Name	Semantics	Sample OPD & OPL	Source	Destination
Consumption link	The process consumes the object.	 **Eating** consumes **Food.**	consumed object	consuming process
Result link	The process generates the object.	 **Mining** yields **Copper.**	creating process	created object
Effect link	The process affects the object by changing it from one state to another state.	 **Purifying** affects **Copper.**	affected object and affecting process are both source and destination	

Name	Semantics	Sample OPD & OPL	Source	Destination
Agent link	Agent is a human or a group of humans who enables the occurrence of the process to which it is linked but is not transformed by that process.	 **Welder** handles **Welding.**	agent – the triggering and enabling object	enabled process
Instrument link	Instrument is an inanimate object that enables the occurrence of the process to which it is linked but is not transformed by that process.	 **Manufacturing** requires **Machine.**	instrument – the enabling object	enabled process

Source: (Dori, 1995), (Dori, Object-Process Methodology - A Holistic System Paradigm, 2002)

APPENDIX E – CASE STUDY 2 MULTI UTILITY SCORES

- **Case Study 2: Lipase Multi-Utility Tradespace scenario 'a'**

RUN#	Total $/50M3	Productivity utility score	Specific activity utility score	Weighted Productivity	Weighted specific activity	Multi-utility score	Productivity weight	0.5
							Specific activity weight	0.5
1	908.49	0.03	-0.38	0.016	-0.192	-0.176		
2	543.30	0.53	-0.25	0.266	-0.125	0.140		
3	178.11	-0.04	-0.43	-0.021	-0.217	-0.238		
4	251.94	-0.12	-0.48	-0.058	-0.240	-0.298		
5	1170.06	0.03	-0.36	0.017	-0.179	-0.162		
6	543.30	0.19	-0.33	0.096	-0.165	-0.069		
7	247.31	0.22	-0.36	0.112	-0.181	-0.069		
8	543.30	0.19	-0.33	0.096	-0.165	-0.069		
9	1040.28	0.52	-0.33	0.262	-0.165	0.097		
10	982.32	-0.02	-0.42	-0.010	-0.209	-0.219		
11	41.94	-0.08	-0.47	-0.038	-0.234	-0.273		
12	37.31	0.22	-0.40	0.109	-0.202	-0.092		
13	562.35	0.00	-0.33	0.000	-0.167	-0.167		
14	438.30	0.13	-0.36	0.065	-0.179	-0.114		
15	648.30	0.15	-0.36	0.074	-0.178	-0.103		
16	524.25	0.40	-0.27	0.202	-0.134	0.068		
17	543.30	0.19	-0.33	0.096	-0.165	-0.069		
18	543.30	0.19	-0.33	0.096	-0.165	-0.069		
19	543.30	0.19	-0.33	0.096	-0.165	-0.069		
20	977.70	0.43	-0.35	0.214	-0.177	0.037		
21	830.28	0.28	-0.34	0.138	-0.172	-0.034		
22	309.89	0.48	-0.35	0.240	-0.176	0.064		
23	509.82	0.03	-0.40	0.017	-0.201	-0.184		
24	772.32	0.03	-0.38	0.014	-0.188	-0.174		
25	99.89	0.36	-0.36	0.178	-0.179	-0.001		
26	439.68	-0.06	-0.45	-0.031	-0.225	-0.257		
27	634.98	0.13	-0.38	0.067	-0.189	-0.122		
28	960.06	0.04	-0.38	0.021	-0.192	-0.171		
29	830.28	0.51	-0.35	0.256	-0.174	0.082		
30	229.68	-0.03	-0.44	-0.016	-0.218	-0.234		
31	543.30	0.48	-0.25	0.242	-0.126	0.117		
32	543.30	0.19	-0.33	0.096	-0.165	-0.069		

- **Case Study 2: Lipase Multi-Utility Tradespace scenario 'b'**

RUN#	Total $/50M3	Productivity utility score	Specific activity utility score	Weighted Productivity	Weighted specific activity	Multi-utility score	Productivity weight	0.2
							Specific activity weight	0.8
1	908.49	0.03	-0.38	0.006	-0.307	-0.301		
2	543.30	0.53	-0.25	0.106	-0.200	-0.094		
3	178.11	-0.04	-0.43	-0.008	-0.347	-0.356		
4	251.94	-0.12	-0.48	-0.023	-0.384	-0.407		
5	1170.06	0.03	-0.36	0.007	-0.287	-0.280		
6	543.30	0.19	-0.33	0.038	-0.264	-0.226		
7	247.31	0.22	-0.36	0.045	-0.290	-0.245		
8	543.30	0.19	-0.33	0.038	-0.264	-0.226		
9	1040.28	0.52	-0.33	0.105	-0.264	-0.159		
10	982.32	-0.02	-0.42	-0.004	-0.334	-0.338		
11	41.94	-0.08	-0.47	-0.015	-0.375	-0.390		
12	37.31	0.22	-0.40	0.044	-0.323	-0.279		
13	562.35	0.00	-0.33	0.000	-0.267	-0.267		
14	438.30	0.13	-0.36	0.026	-0.286	-0.260		
15	648.30	0.15	-0.36	0.030	-0.285	-0.255		
16	524.25	0.40	-0.27	0.081	-0.215	-0.134		
17	543.30	0.19	-0.33	0.038	-0.264	-0.226		
18	543.30	0.19	-0.33	0.038	-0.264	-0.226		
19	543.30	0.19	-0.33	0.038	-0.264	-0.226		
20	977.70	0.43	-0.35	0.086	-0.284	-0.198		
21	830.28	0.28	-0.34	0.055	-0.275	-0.220		
22	309.89	0.48	-0.35	0.096	-0.281	-0.186		
23	509.82	0.03	-0.40	0.007	-0.322	-0.315		
24	772.32	0.03	-0.38	0.006	-0.301	-0.295		
25	99.89	0.36	-0.36	0.071	-0.287	-0.216		
26	439.68	-0.06	-0.45	-0.012	-0.361	-0.373		
27	634.98	0.13	-0.38	0.027	-0.303	-0.276		
28	960.06	0.04	-0.38	0.009	-0.308	-0.299		
29	830.28	0.51	-0.35	0.103	-0.278	-0.176		
30	229.68	-0.03	-0.44	-0.006	-0.348	-0.355		
31	543.30	0.48	-0.25	0.097	-0.201	-0.104		
32	543.30	0.19	-0.33	0.038	-0.264	-0.226		

- **Case Study 2: Lipase Multi-Utility Tradespace scenario 'c'**

RUN#	Total $/50M3	Productivity utility score	Specific activity utility score	Weighted Productivity	Weighted specific activity	Multi-utility score	Productivity weight	0.8
							Specific activity weight	0.2
1	908.49	0.03	-0.38	0.026	-0.077	-0.051		
2	543.30	0.53	-0.25	0.425	-0.050	0.375		
3	178.11	-0.04	-0.43	-0.033	-0.087	-0.120		
4	251.94	-0.12	-0.48	-0.093	-0.096	-0.189		
5	1170.06	0.03	-0.36	0.027	-0.072	-0.044		
6	543.30	0.19	-0.33	0.153	-0.066	0.087		
7	247.31	0.22	-0.36	0.180	-0.073	0.107		
8	543.30	0.19	-0.33	0.153	-0.066	0.087		
9	1040.28	0.52	-0.33	0.419	-0.066	0.353		
10	982.32	-0.02	-0.42	-0.017	-0.083	-0.100		
11	41.94	-0.08	-0.47	-0.061	-0.094	-0.155		
12	37.31	0.22	-0.40	0.175	-0.081	0.094		
13	562.35	0.00	-0.33	0.000	-0.067	-0.067		
14	438.30	0.13	-0.36	0.104	-0.072	0.033		
15	648.30	0.15	-0.36	0.119	-0.071	0.048		
16	524.25	0.40	-0.27	0.324	-0.054	0.270		
17	543.30	0.19	-0.33	0.153	-0.066	0.087		
18	543.30	0.19	-0.33	0.153	-0.066	0.087		
19	543.30	0.19	-0.33	0.153	-0.066	0.087		
20	977.70	0.43	-0.35	0.343	-0.071	0.272		
21	830.28	0.28	-0.34	0.221	-0.069	0.152		
22	309.89	0.48	-0.35	0.383	-0.070	0.313		
23	509.82	0.03	-0.40	0.027	-0.080	-0.054		
24	772.32	0.03	-0.38	0.023	-0.075	-0.052		
25	99.89	0.36	-0.36	0.285	-0.072	0.214		
26	439.68	-0.06	-0.45	-0.050	-0.090	-0.140		
27	634.98	0.13	-0.38	0.108	-0.076	0.032		
28	960.06	0.04	-0.38	0.034	-0.077	-0.043		
29	830.28	0.51	-0.35	0.410	-0.070	0.340		
30	229.68	-0.03	-0.44	-0.026	-0.087	-0.113		
31	543.30	0.48	-0.25	0.388	-0.050	0.337		
32	543.30	0.19	-0.33	0.153	-0.066	0.087		