DTU Library

An Integrated Methodology for Emulsified Formulated Product Design

Mattei, Michele

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Mattei, M. (2014). *An Integrated Methodology for Emulsified Formulated Product Design.* Technical University of Denmark, Department of Chemical and Biochemical Engineering.

Advances in Chemical Product Design

An Integrated Methodology for
Emulsified Formulated Product Design

PhD Thesis

Michele Mattei

July 2014

Computer Aided Process and Product Engineering Center

Department of Chemical and Biochemical Engineering

Technical University of Denmark

Preface

This thesis is submitted as partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) at the Technical University of Denmark (DTU).

The work has been carried out at the Computer Aided Process-Product Engineering Center (CAPEC) of the Department of Chemical and Biochemical Engineering, from August 2011 to July 2014, under the supervision of Professor Rafiqul Gani and Professor Georgios M. Kontogeorgis, from the Center for Energy Resources Engineering (CERE) at the Department of Chemical and Biochemical Engineering. The project included also an external research stay period, from January to March 2014, to the Hong Kong University of Science and Technology (HKUST), under the supervision of Professor Ka M. Ng, from the Department of Chemical and Biomolecular Engineering.

I would like to acknowledge, first of all, my supervisors Rafiqul Gani and Georgios M. Kontogeorgis for their teaching, their training, their guidance and their support during these 3 years. I am sure that the time I experienced working with them has deeply changed and improved my personal and technical skills, and I will always be grateful to them for this.

I also would like to genuinely thank Ka M. Ng for having welcomed and hosted me in his group, during my external stay at the HKUST, and for having shared with me his own perspective on my research work, that has opened my eyes and broadened my perspectives. Together with him, I would like to thank Dr. Kelvin Fung, Fanny, Kee, Grace and Prudence for their fundamental help in the laboratory, as well as for their company and support during my stay in Hong Kong.

I also owe thanks to those people I have been collaborating during my project, that is: Dr. Gordon Bell from Syngenta (United Kingdom), Professor Michael Hill from Columbia University (USA), Professor Romain Privat and Professor Jean-Noel Jaubert from the University of Lorraine (France), and Bo Depner and Peter Krogh from Navadan (Denmark). Each of them contributed to my project by providing new challenges, and/or helping me to solve actual ones.

I cannot forget to thank every single Master student, PhD student, Postdoc and faculty at CAPEC and PROCESS for collaborating to create a wonderful, happy atmosphere in the research center, making everything easier, in the good as well as in the bad times.

Finally, I would like to thank my family, my friends and my girl-friend, for having been close to me, virtually from Italy, as well as physically in Copenhagen, during this 3 long years. I would have never make it without them.

..to my mum..

Abstract

The consumer oriented chemical based products are used every day by millions of people. They are structured products constituted of numerous chemicals, and many of them, especially household and personal care products, are emulsions where active ingredients, solvents, additives and surfactants are mixed together to determine the desired emulsified product. They are still mainly designed and analysed through trial-and-error based experimental techniques, therefore a systematic approach, integrating model-based as well as experiment-based techniques, for design of these products could significantly reduce both time and cost connected to product development by doing only the necessary experiments, and ensuring chances for innovation.

The main contribution of this project is the development of an integrated methodology for the design of emulsified formulated products. The methodology consists of three stages: the problem definition stage, the model-based design stage, and the experiment-based verification stage. In the problem definition stage, the consumer needs are translated into a set of target thermo-physical properties and into a list of categories of ingredients that are to be included in the formulation. In the model-based design stage, structured databases, dedicated algorithms and a property model library are employed for designing a candidate base case formulation. Finally, in the experiment-based verification stage, the properties and performances of the proposed formulation are measured by means of tailor-made experiments. The formulation is then validated or, if necessary, refined thanks to a systematic list of action.

The problem definition stage relies on a robust knowledge base, which needs to systematically generate quantitative, useful input information for the model-based stage, starting from the consumer assessments. In the model-based stage, comprehensive chemical databases, consistent property models and a dedicated algorithm for the design of emulsified solvent mixtures are needed. Finally, for the experiment-based stage, an efficient planning of the experiments is required, together with the systematic generation of a list of actions to be taken, in case some of the experiments do not validate the candidate formulation generated in the previous stage.

All the above mentioned issues are addressed in this PhD work: the necessary property models have been retrieved and organized in a model library; new property models have been developed for a set of thermo-physical properties of surfactants; a robust, systematic knowledge-base has been developed in relation to emulsified formulated products; chemical databases have been improved and generated; and an algorithm for the model-based design of emulsified solvent mixtures has been developed. All these tools have been implemented as a new template in the virtual Product-Process Design laboratory software. To illustrate the application of the proposed methodology, three case studies have been developed. For one of these case studies, the whole methodology has been applied, while for the other two, only the first two stages and part of the experiment-based verification

stage have been applied, that is, the experimental work has been planned, a list of actions has been generated, but no actual measurement has been taken.

Resumé på dansk

Forbrugerorienterede, kemisk baserede produkter anvendes dagligt af millioner af mennesker. De er strukturerede produkter bestående af en lang række kemikalier. Mange og især produkter til husholdning og personlig pleje er emulsioner, hvor aktive ingredienser, opløsningsmidler, tilsætningsstoffer og tensider er blandet sammen for at opnå det ønskede emulgerede produkt. De er stadig primært designet og analyseret via "trial-and-error"-baserede eksperimentelle teknikker. Derfor vil en systematisk tilgang, der integrerer model- og eksperimentbaserede teknikker til design af disse produkter kunne reducere både tid og omkostninger forbundet med produktudvikling, idet kun de nødvendige eksperimenter bliver udført, og åbner desuden mulighed for at fremme innovation.

Det vigtigste bidrag i dette projekt er udviklingen af en integreret metode til design af emulgerede, formulerede produkter. Metoden er opdelt i tre faser: Problemformulering, modelbaseret projektfase, og forsøgsbaseret verifikationsfase. Under problemformuleringen, omsættes forbrugernes behov til et sæt af ønskede termo-fysiske egenskaber og til en liste over kategorier af ingredienser, som skal indgå i formuleringen. I den modelbaserede projektfase anvendes strukturerede databaser, dedikerede algoritmer og et egenskabsmodelbibliotek til at designe en base-case kandidatformulering. Endelig, i den eksperimentbaserede verifikationsfase, afmåles egenskaber og funktioner af den foreslåede formulering ved hjælp af skræddersyede eksperimenter. Derefter valideres formuleringen eller, om nødvendigt, raffineres via en systematisk fremgangsmåde.

Problemformuleringsfasen bygger på en solid vidensbase, som systematisk skulle kunne generere kvantitative, nyttige oplysninger som input til det modelbaserede fase, startende fra forbrugernes vurderinger. I den modelbaserede fase er der behov for omfattende kemiske databaser, konsistente egenskabsmodeller og en dedikeret algoritme til design af emulgerede opløsningsmiddelblandinger. Til den eksperimentbaserede fase kræves en effektiv planlægning af forsøg samt en systematisk generering af en liste over nødvendige tiltag, i tilfældet at kandidatformulering ikke kan valideres eksperimentelt.

Alle de ovennævnte udfordringer er behandlet i dette ph.d.-projekt: De nødvendige egenskabsmodeller er blevet identificeret og organiseret i et modelbibliotek; der er blevet udviklet nye egenskabsmodeller for et sæt af termo-fysiske egenskaber for overfladeaktive stoffer; en robust, systematisk vidensbase er udviklet over emulgerede, formulerede produkter; kemiske databaser er blevet forbedret og genereret; og en algoritme til modelbaseret design af emulgerede opløsningsmiddelblandinger er blevet udviklet. Alle disse værktøjer er blevet implementeret som en ny skabelon i det virtuelle Product-Process Design Laboratorium software.

For at illustrere den udviklede metode er tre casestudier blevet formuleret. Metoden er anvendt i sit fulde omfang på ét casestudie, mens metodens første to faser samt en del af den forsøgsbaserede verifikationsfase er anvendt på de to resterende casestudier. Med

sidstnævnte menes en liste over eksperimentelle tiltag er blevet genereret, men der er ikke skredet til yderligere handling.

Table of Contents

ix

1

INTRODUCTION

Reklaitis (2014) observed that process systems engineering, which in content if not in title has a history almost as old as chemical engineering itself, has had a remarkable impact on the chemical and related industry. Robust and reliable model based methodologies have been deeply applied in the fields of process design, process control and operations, and they are now employed as common practice throughout the industry world-wide (Stephanopoulos and Reklaitis, 2011). It has been only in the 1990's, however, that the interest in chemical product design and engineering emerged, stimulated by the contributions to areas such as the selection and design of solvent systems and working fluids.

Background in transport phenomena and thermodynamics have been applied to understand the processes characterizing the properties and performances of chemical products. The process systems engineering community contributed by introducing mathematical modeling, especially in the application where the product properties and performances can be related to the molecular structure of chemicals; an excellent overview of such models has been given by Achenie et al. (2003). Moreover, the process systems engineering community satisfied the necessity of capturing the knowledge of product designers, storing it and then providing it in terms of explicit knowledge to be manipulated to perform predictions that can be used for the design and development of innovative products (Joglekar, Giridhar and Reklaitis, 2014).

Gani (2004) emphasized the complexity of chemical product design with the concept of the chemical product tree, shown in Figure 1.1. It gives an idea of the size of the recent shift that is observed in the chemical industry from materials valued for their purity, such as, the commodities, to materials sold for their performance behavior, such as consumer products (Villadsen, 1997).

Refined Chemicals & Consumer Products (≈ 30000)
Plastics, pharmaceuticals, dyes, solvents, fertilizers, fibres, dispensers, cosmetics,

Intermediate Products (≈ 300)
Methanol, vinyl chloride, styrene, urea, formaldehyde, ethylene oxide, acetic acid, acrylonitrile, cyclohexane, acrylic acid

Basic Products (≈ 20)
Ethylene, propene, butadiene, benzene, synthesis-gas, actylene, ammonia, sulfuric acid, sodium hydroxide, chlorine

Raw Materials (≈ 10)
Petroleum, natural gas, coal, biomass
Rock, salt, phosphate, sulfur, air, water

Figure 1.1 The chemical product tree: classification
of chemical-based products (from Gani, 2004)

The roots of the tree consist of a limited number of raw materials, which are processed to obtain the commodity products (basic products). Specialty chemicals (intermediate products) are then manufactured from the commodities, and finally the leaves of the tree represent a very large portfolio of higher value products (refined chemicals and consumer products), obtained by processing and/or combining the chemicals of the previous classes. As one ascends the chemical product tree, the number of products belonging to each category grows exponentially from around 10 for the raw materials, up to almost 30,000 in the last class of higher value added products.

This last class of chemical products includes formulations, devices and technology based consumer goods. Formulated products consist of pharmaceuticals, paints, food, cosmetics, detergents, pesticides, *etc.*, in which 5 to more than 20 ingredients are usually present, representing a wide range of chemical compounds, such as, polymers, surfactants, solid particles, solvents, pigments, and aromas (Abildskov and Kontogeorgis, 2004).

Cussler and Moggridge (2011), in a certainly very influential book in the field of chemical product engineering, classify the chemical products into four categories:

- Commodities: like ethylene and ammonia, made in large quantities and sold at the lowest price possible;
- Devices: miniature processes which accomplish a particular chemical transformation at a small scale, as an artificial kidney, or a mosquito repellent vaporizer;
- Molecules: most often exemplified by pharmaceuticals, where the key is not process cost or convenience but discovery of the active compound in the first place;
- Structures: including several consumer products, such as sunscreens and food products, where the key is the product function.

Irrespectively to the product type, the common practice, in the development of such products, is still the experiment-based and trial-and-error approach. However, a systematic integrated procedure, where candidate higher added value products are designed through a model-based methodology, and then validated and/or refined by means of dedicated experiments, represents an efficient alternative, with respect to time and resources, speeding up the product development.

Wintermantel (1999) recognizes that the development of methodologies, tools and strategies is crucial in order to systematically improve the design and analysis of chemical products, as speeding up the product development is of paramount importance (Charpentier, 2009).

Recently, many efforts have been taken to develop systematic methodologies in the product design area, using computer-aided tools. A non-exhaustive list is given below:

- Molecular design (Harper and Gani, 2000);
- Solvent design (Gani, 2005);
- Mixture design (Eden et al., 2004);
- Polymer design (Satyanarayana et al., 2009);

Even more recently, various attempts have been made to tackle the design of formulated products, by means of model-based methodologies. The most remarkable have been proposed by Conte, Gani and Ng (2011), in relation to homogeneous formulated products, and by Yunus et al. (2014), focusing on tailor-made blended products. These methodologies are based on the "define target – match target" paradigm, employing the reverse design techniques. That is, the consumer needs of the product are the known variables and they are the input of the methodology; these are then converted into a defined set of target properties, driving the selection and design of candidate formulation ingredients by means of appropriate property models. At the same time, the formulation composition that satisfy the product constraints is determined, using suitable mixture property models, as well as phase stability algorithm.

1.1 Motivation and Problem Definition

Many, among the consumer chemical-based products, are the physical forms of the formulations: suspension containing insoluble chemicals dispersed in the liquid mixture with the help of proper dispersing agents; emulsions where immiscible liquids are kept together through selected emulsifying agents; solid products such as pharmaceutical tablets or soap bars. The performances of such structured products are related not only to the presence of active ingredients and additives in the formulation, but also to the product's structural and material properties (Smith and Ierapepritou, 2010).

Among this category of products, emulsified formulated products are the most relevant, with particular application in the food and cosmetic industries (Mattei, Kontogeorgis and Gani, 2014). Emulsions are defined as mixtures of two normally immiscible liquids, kinetically stabilized by emulsifying agents (most often surface active agents, that is, surfactants) that lie at the interface between the two liquid phases. Active ingredients and

additives are then dissolved in the continuous and/or in the dispersed phases, according to the needs of the product.

With regards to emulsified formulated products, Bernardo and Saraiva (2012) proposed an approach to simultaneously tackle product and process design, with particular attention to cosmetic emulsions. Bagajewicz et al. (2011), instead, proposed an approach integrating the concept of price-competitive markets. However, there is a lack of works where process systems engineering tools are applied for the solution of problems involving the design of emulsified formulated products.

In this work, a systematic methodology for the design of emulsified formulated products is proposed. The integrated methodology consists of three stages: a problem definition stage where the consumer needs are converted into valuable, quantitative information; a model-based stage where property models, structured databases and dedicated algorithms are applied together in order to propose a candidate formulated products; and finally an experiment-based stage for validation and/or refinement of the candidate product, thereby reaching the final emulsified formulation. Each of the stages above is divided into tasks, and every task is constituted of several sub-tasks, in which different methods and tools are employed. Systematic data-flow and work-flow link together the different sub-tasks of the framework, generating a systematic methodology leading to the final product recipe.

The necessary models integrated in the methodology have been retrieved in the literature from various sources and, when they were not available, they have been developed. A systematic, step-by-step algorithm for the design of solvent mixtures in the emulsified form (EMUD) has also been developed, and integrated with other algorithm previously developed by Conte et al. (2010) for the design of homogeneous formulations (MIXD and STABILITY), forming a robust tool for mixture design. Databases of active ingredients as well as additives have been built, filled with all the properties that are necessary in the design procedure. Moreover, a systematic knowledge-base has been developed, for an easy and reliable solution of the problem definition stage. The overall methodology has been highlighted through three case studies, consisting of a UV sunscreen, a tank-cleaning detergent and a hand-wash, all in the emulsified form. Finally, the overall work-flow, the models, the methods, and the tools, have been integrated as a new template into an in-house software, the virtual Product-Process Design Laboratory, for the design and analysis of emulsified formulated products.

1.2 Structure of the Thesis

This PhD thesis is divided into seven chapters. The current chapter (Chapter 1) briefly introduces the concept of chemical product design, and underlines the motivation for focusing the attention in this relatively new area of chemical engineering, and in particular in the area of emulsified formulated products.

Chapter 2 gives the theoretical background of chemical product design, providing a detailed explanation of the objectives of this PhD work. Here, not only fundamental concepts of chemical product design are introduced, but also the current state of the art and perspectives are given. In addition, the concept of surfactants and emulsions is introduced, and the issues and needs are underlined, giving the basis of the contents of the following chapters.

Chapter 3 is dedicated to the property modeling in emulsified formulated product design. First, the general property models satisfying the needs of formulation design are presented, divided between pure component and mixture property models. Then, those property models that are specific of surfactant and/or emulsions are described, and the modeling efforts that have been made during this PhD work for developing/improving property models for emulsified formulated product design are highlighted.

Chapter 4 highlights the methods and tools specially developed in this work. A systematic knowledge-base for storing and providing, when needed, in an efficient way the necessary information and data for product development and design is presented first. Then the structured databases that have been adopted and/or built in this work are described, and finally an algorithm specifically developed for the design of solvent mixtures in the emulsified form (EMUD) based on the reverse approach is presented.

In Chapter 5, the integrated methodology for the design of emulsified formulated products is illustrated. Details of the different tasks and sub-tasks constituting the methodology are given, together with the work-flow, the data-flow and the necessary tools and methods. A brief introduction to the new template integrated into the virtual Product-Process Design Laboratory is also given here.

Chapter 6 presents the case studies that have been developed in this work: a UV sunscreen, a tank-cleaning detergent and a hand-wash, where only for the third case-study, the complete methodology has been applied. In relation to the first two case studies, in fact, no experimental work has been made.

Chapter 7 is the conclusive chapter. Here the achievements obtained in this PhD work are summarized, and the challenges and future perspectives in the field of chemical product design are given.

2

PRODUCT DESIGN: AN OVERVIEW

This chapter introduces general definitions and classifications commonly used in the chemical product design, which are adopted in the next chapter. In §2.1, the state of the art of chemical process design is given, followed by the main concepts of surfactants and emulsions and the consequences of their inclusion in the area of chemical product design (§2.2). In §2.3 the issues and needs related to the design of emulsified formulations are presented, and finally, in §2.4, the contribution of this work to the chemical product design community is highlighted.

2.1 State of the Art

Chemical product design is an extremely wide area: an amine for scrubbing acid gases, a pollution-preventing ink, an electrode separator for high power batteries, and a ventilator for a well-insulated house are all example of chemical products, according to Moggridge and Cussler (2000). There seems to be nothing in common between these products, but in fact the procedure by which they are designed is the same. Moggridge and Cussler (2000) propose a four-level conceptual procedure for chemical-product design:

- Define the consumer needs;
- Generate ideas to fulfill these requirements;
- Select the most advantageous option among those generated;
- Manufacture the product.

The chemical product of interest can be a device, a single chemical, or a mixture/blend. Chemical products as mixtures and blends are usually considered when a single molecule is unable to provide all the desired product functions, or perform all the desired process

tasks. Solvent mixtures, tailor-made fuels, lubricants are examples of blended chemical products.

A wide variety of computer-aided methods and tools have been developed for design of chemical products represented by the properties of a single molecules. These methods are classified under CAMD (Computer Aided Molecular Design) and they follow the same main steps: generate feasible chemical structures, estimate the thermo-physical properties through property models, and select the molecules that match the desired targets. The main features for these methods and tools are that building-blocks based methods are used to represent the molecular structures and their contributions are used to estimate the target properties (Gani and Ng, 2014).

For the computer-aided mixture-blend design (CAM^{b}D), the properties and algorithms are different, even though similar building-block based property models may be used. Here, the properties depend on the mixture composition, and the design algorithm needs to identify the molecules and their compositions in solution matching the target properties. Systematic decomposition based solution approaches are usually employed to manage the complexity of these design problems efficiently, by reducing the search space in subsequent steps (Yunus et al., 2014).

Many consumer oriented chemical based products, however, are formulations. These may contain materials from different classes of chemicals, such as polymers, surfactants, solvents, pigments and aromas. Conte et al. (2011) classify these classes of chemical as follows:

- Active ingredients: these chemicals are the most important ones in the formulation, because they satisfy the main needs of the product, thus defining the function of the product itself;
- Solvent mixture: it is usually present in high concentration in the formulation and it has the function of dissolving the active ingredients and other chemicals in the formulation, ensuring the product to be in the desired physical form, and to be property delivered;
- Additives: these chemicals are usually present in low concentration and they satisfy the secondary needs of the product, enhancing the end-use product properties.

In terms of solution approaches to product design, Ng, Gani and Dam-Johansen (2007) identify three types:

- Experiment-base trial-and-error approach: this approach is employed when mathematical models for the estimation of target properties are not available. A large number of consumer products are currently developed through this approach, where past knowledge and experience are crucial;
- Model-based approach: when validated mathematical models for the estimation of the target properties are available, a list of feasible candidates is efficiently and quickly generated and tested. CAMD and CAM^{b}D enter in this type of solution approach to product design;

- Integrated experiment-modelling approach: this approach is used when mathematical models are not available for all the target properties. The design problem is decomposed into a hierarchical sequence of sub-problems: as one goes from the outer levels to the inner levels, the number of candidate decreases and the inner levels employ correlations and/or experiments for the final product validation.

Table 2.1 illustrates the characteristics of the above mentioned solution approaches.

Table 2.1 Comparison between different solution approaches
to chemical product design

	Experiment-based Approach	Model-based Approach	Integrated Approach
Tools	Experimental setup	Mathematical models	Mathematical models and experimental setup
Development time	Long	Short	Short
Consumption of resources	High	None	Limited
Uncertainty	Limited	High	Limited
Need for validation	No	Yes	No
Chance for innovation	Limited	High	High

2.2 Emulsion-based Chemical Product Design

In this paragraph, the fundamental concepts and properties of surfactants and emulsions are presented and their influence on emulsified formulated product design is underlined. Finally, the issues and needs to be faced in relation to the design of such products are discussed.

2.2.1 Surfactants

Surface-active agents, or simply surfactants, can be found in almost any household and personal care product, as well as in nearly every industrial process. They are chemicals showing ability to adsorb at interfaces, either between solid and liquid (suspensions), between liquid and gas (foams) and between two liquids (emulsions). Despite surfactants are often present in very small amounts in the products, they greatly affect the overall properties of the systems, because of their ability (Porter, 1994). For this reasons, many surfactants are sometimes referred to as effect chemicals, such as, dispersing agents, foaming agents, emulsifying agents, instead of using the general term surfactants.

Surfactants are chemicals having an amphiphilic nature, as their molecular structures contain two part with different polarity: a non-polar, hydrophobic part (usually referred to as

the hydrophobic tail), and a polar, hydrophilic part (the hydrophilic head). The hydrophobic part is commonly a hydrocarbon which may contain branches and/or aromatic structure, while the hydrophilic part is an ionic or strongly polar group.

A surfactant molecule is not fully compatible with either a non-polar or polar medium: there is always a conflict between the affinity of the head-group and the tail, and this gives surfactants their unique properties (Landgren, 1990). The chemical structure plays therefore a fundamental role for the thermo-physical properties of the surfactants.

Surfactants are classified according to the nature of their head group. Four main types of surfactants are identified: anionic, cationic, non-ionic and amphoteric surfactants. Anionic and non-ionic surfactants account for more than 80% of the manufactured surfactants (Pilemand, 2002). Figure 2.1 schematically represents the different surfactant types.
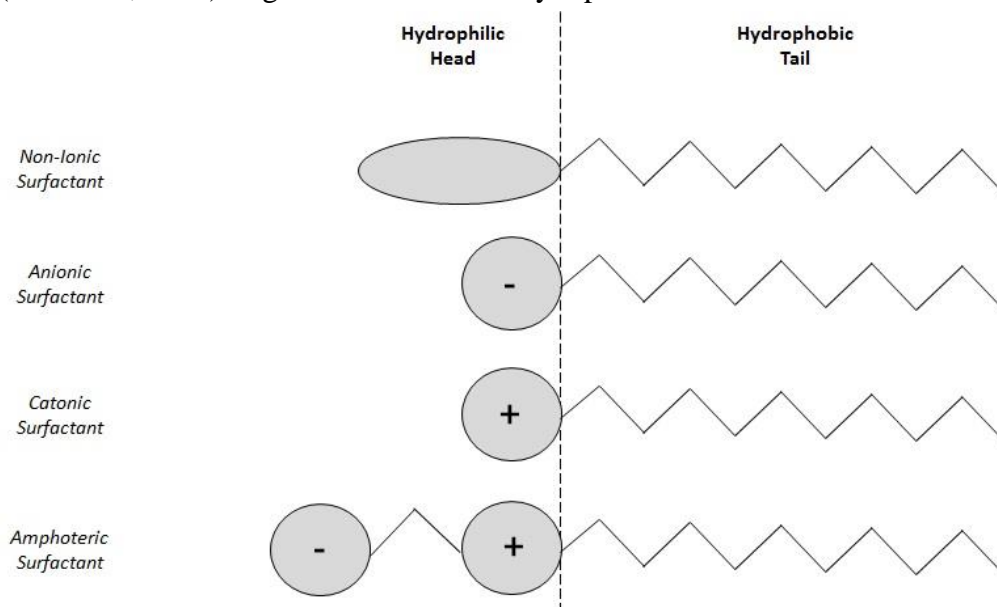
Figure 2.1 Surfactant classification according to the nature of the head group

The ability of the surfactant to adsorb at interfaces and to form aggregates called micelles, determines the properties of the surfactant. It is adsorption at interfaces that gives the surface-active effects of foaming, wetting, emulsification, dispersion and detergency, whereas the micellar properties give the surfactant solution its bulk properties such as, for example, the viscosity. Properties such as the critical micelle concentration (CMC), the Krafft temperature (T_K), the cloud point (CP), the surface tension (σ), the hydrophilic-lipophilic balance (HLB) and the phase inversion temperature (PIT) are fundamental to predict the effects of the addition of surfactants in solution.

The critical micelle concentration of a surfactant indicates the point at which monolayer adsorption of surfactant molecule at the interface is complete; further addition of surfactant molecules causes them to aggregate into micelles. Since micelles themselves are not surface-active, the surface tension remains approximately constant beyond the CMC. There is a considerable interest in the CMC property, as this represents, in practice, the

lowest surfactant concentration needed to get maximum benefit if surfactant abilities based on surface-activity are needed (e.g. wetting and foaming).

The Krafft temperature, or Krafft point, is the temperature below which the surfactant molecules do not associate into aggregates. Such a behavior is mostly observed for ionic surfactants, as only few non-ionic surfactants possess a Krafft temperature. Thus, a formulation can be manufactured in the form of an emulsion only above the Krafft temperature of the surfactant system.

Non-ionic surfactants, instead, show a miscibility gap, in aqueous solution, with an upper critical temperature. This temperature is called the cloud point, as the mixture separate into two phases at higher temperatures, becoming cloudy. The cloud point is characteristic of non-ionic surfactant containing polyethylene oxide chains in their head groups. Cloud point temperatures are an excellent means of screening candidate surfactants, as stable emulsified formulated products can be obtained only below the cloud point of the surfactant system.

The surface tension is a measure of the wetting ability of a compound and it is a very important property for surfactants. As the surfactant molecules adsorb at the interface, in an aqueous solution, the surface tension of the mixture decreases until reaching a minimum in correspondence to the critical micelle concentration. In emulsified formulated product design, as surfactants are added in higher concentrations than their CMC, the surface tension of aqueous surfactant mixtures are determined by the choice of the surfactant system.

The hydrophilic-lipophilic balance system is probably the most common method to correlate the surfactant structure with the effectiveness of the surfactant as emulsifiers. As the Bancroft rule states that the phase in which a surfactant is more soluble constitutes the continuous phase (Bancroft, 1915), the HLB value indicates, on an arbitrary scale, how the surfactant will behave in a solution with an aqueous and an organic phase. HLB values below 10 indicates the formation of a water-in-oil emulsion, while surfactants with HLB values higher than 10 preferentially yield to oil-in-water emulsions. The HLB values are often used as an initial guide to screen candidate surfactants, as the application of this system is limited by the fact that the emulsion type is frequently affected also by the water-oil ratio, the temperature, *etc*.

An alternative method for characterizing surfactants is the phase inversion temperature. This is defined as the temperature at which the emulsion changes from an oil-in-water to a water-in-oil emulsion (Duncan, 1983). For application in the product design area, an emulsified formulation has to be far enough from its phase inversion temperature, not to risk to change its physical form, thus altering the desired properties.

2.2.2 Emulsions

Emulsions are defined as dispersed systems for which the phases are immiscible or partially miscible liquids. Emulsions are dispersions of one liquid in another, typically classified as oil-in-water or water-in-oil emulsion, depending if the continuous phase is the

aqueous or the organic phase, respectively. They are characterized by droplet size of about 1 μm, and they are typically unstable systems, which will eventually separate and require emulsifiers, most often surfactants, to be kinetically stabilized.

Emulsions find extensive applications in the food, pharmaceutical, and cosmetic industries. They are highly complex multicomponent systems, containing surfactants, solvents, preservatives, and numerous other compounds.

A very important issue in emulsion design is to define/design the right type of emulsion for a certain application. The emulsion type can be determined from the HLB system, or from the phase inversion temperature, as explained in §2.2.1. In addition, the performances of an emulsified product are related to the presence of active ingredients and additives, but also to the product's structural and material properties (Smith and Ierapepritou, 2010).

2.3 Issues and Needs

Given the current state of the art of chemical product design area (§2.1), and the new concepts introduced with the design of emulsified formulations (§2.2), the research issues and needs to be faced are many and diverse. They can be organized under the following generic points:

- Problem definition;
- Property models;
- Methods and tools;
- Methodologies;
- Systematic frameworks;
- Multidisciplinary modelling;
- Multiscale modelling.

Problem definition

The reliability of a solution to a generic product design problem relies on the problem definition. This stage is responsible of the identification of the consumer needs for a specific products, and of relating these needs to a set of desired target properties. There is the need for the development of a reliable knowledge base that may guide the product designer to convert the problem from consumer assessments to quantitative technical specifications, such as, thermo-physical properties (Harper, 2000). This is very relevant also for identifying the relationship between product performances, product composition, ingredients properties, *etc.*

Property models

Models are the core of all computer-aided product design problems, since the reliability of the solution largely depends on the choice of the models and their uncertainties. The biggest challenge to the use of property models is the scarce availability of model parameters; if model parameters are not available for a candidate ingredient of the formulation,

this molecule has to be discarded, since its properties cannot be estimated. This way, a potential optimal candidate is rejected.

The major need in this area is to extend the application range of existing property models, improving their performances, and, if necessary, develop new reliable property models. Most of the models required for the design of emulsified formulated products, in fact, are not currently available in a form that can be implemented as a part of a computer-aided product design method, and the model parameters are restricted only to a few systems.

Methods and tools

It is necessary to develop design algorithms the focus of which is not only the product cost, but also product and process performances, environmental impact, and health and safety concerns over the entire product life cycle. These approaches have to be systematic but flexible, simple but accurate, and they should be able to solve these problems with the reverse approach (Gani and Pistikopoulos, 2002).

Databases of chemicals are also to be extended and/or built, in order to consider a wider range of candidate ingredients to be screened on the basis of property models and design algorithms.

Methodologies

Costly and time consuming trial-and-error experimental procedures are usually performed for the design of chemical product design. The development of systematic procedures, with related work-flows and data-flows, where computer-aided tools are employed for a first screening of thousands of candidate, saving the valuable experimental resources for focused experiments, has been recognized as one of the main research challenges in the context of chemical product engineering (Conte et al., 2011).

Systematic frameworks

The solution of a chemical product design problem requires different methods and tools, such as, structured databases, property prediction models, design algorithms and many more steps. The development of such methods and tools is almost as important as the integration of them into a systematic framework that allows inter-changes of information, data and results. The structure of such a framework should be flexible, so that new models can added, and new classes of products can be designed by using a common work-flow. In addition, a user-friendly interface is required, so that the software can be employed for industrial application, as well as for effective teaching of chemical product engineering.

Multidisciplinary modelling

The product design framework should take into account not only the product structure and composition, but also the manufacturing investments and costs, the associated supply chain and marketing issues, such as the consumer behavior with respect to the product price (Bagajewicz, 2007). Integrative approaches, involving marketing and management

issues on the business side, and product design and prototyping on the technical side, are necessary for the development of chemical-based products (Cheng et al., 2009). Multi-disciplinary approaches need to be developed to address the increasing attention to environmental, safety and social requirements, and to the transition towards sustainability (Charpentier and McKenna, 2004). Considering such a multidisciplinarity would return the product design closer to practice in industry.

Multiscale modelling

It is necessary to consider different complexity levels in chemical product and process engineering, in order to understand and describe phenomena at different scales, and being able to identify and model the relationships between them. The understanding of product behaviors at nano- and micro-scales is still largely incomplete, and the relationship between the manufacturing protocol of structured products (such as, emulsified formulations) and their internal micro-structure are mainly described by means of heuristics. A systematic methodology for simultaneous product and process design is currently regarded as one of the major challenges to be face by the chemical process and product engineering (Martin and Martinez, 2013).

2.4 Addressing the Complexity

This PhD work addresses some of the issues and needs that have been highlighted in §2.3.

Property models (Chapter 3)

When considering consumer oriented chemical based products, the performances of the products are the key for the design of successful formulations. They are influenced by a wider range of thermo-physical properties of the ingredients as well as of the mixture. Consequently, property models are considered the core of chemical product design (Kontogeorgis and Gani, 2004). Pure component thermo-physical properties are necessary, at first, as they are the basis for the estimation of the mixture properties. Solubility and miscibility issues are also a major concern, therefore phase equilibria related models are to be employed. In this work, existing models for the estimation of target pure component and mixture thermo-physical properties have to be adopted and, in case they are not available of lacking of accuracy, new property models need to be developed.

Knowledge base (Chapter 4)

Several information are necessary for supporting the decision and choices required during the design of a consumer oriented chemical based products. The consumer needs are to be identified and translated into target properties and numerical boundaries and, in case quantitative models for the relative target properties are not available, categories of chemicals able to satisfy the needs have to be defined. In this work, all the type of information which could be useful in the design and verification of emulsified formulated products are collected and stored in the knowledge base, for easy retrieval and use when necessary.

Structured databases (Chapter 4)

Databases of chemicals are required for screening and selecting the most suitable chemicals to be included in the formulation, on the basis of the target properties. Formulated products often consists of 5 to 20 different ingredients, each providing a specific needed performance; therefore, databases containing all the candidate ingredients classified depending on their role and function in the formulation are necessary. In this work, databases have been collected, improved, and built, according to the needs set by the case studies considered for the design of emulsified formulated products.

Algorithm (Chapter 4)

Computer-aided mixture design is still considered quite an immature area (Gani, 2004), and there is just limited knowledge and know-how about systematic approaches for the design of this type of chemical products. Recent efforts have been directed to the design of liquid solvent mixtures (Klein et al., 1992; Karunanithi et al., 2005; Conte et al., 2011) but very little has been done in the area of emulsions. In this work, an algorithm for the design of solvent mixtures in the emulsified form (EMUD) has been developed. This algorithm decomposes the computer-aided mixture design problem into a series of sub-problems of increasing complexity, according to the type of property models used: pure component property models, linear mixture property models, non-linear mixture property models, stability models. Mathematical programming techniques are applied in each of the sub-problems in order to reduce the number of feasible emulsified solvent mixtures.

Methodology (Chapter 5)

A systematic methodology integrating model-based and experiment-based techniques for the efficient design of emulsified formulated products is to be developed. The complexity of the chemical systems to be designed is managed by decomposing them into sub-problems, each of which employing different tools (property models, knowledge base, databases, solvent mixture design algorithms, *etc.*) for its solution.

Framework (Chapter 5)

The methodology, together with the related models, methods and tools are to be collected in a systematic framework so that their use is made more efficient for the design of emulsified formulated products. The framework has to include all the sub-problems relative to the computer-aided stages of the methodology, and it has to be implemented into the virtual Product-Process Design Laboratory software, that already hosts a general work-flow and specific templates for the solution of product and process design problems.

Case Studies (Chapter 6)

The application of the methodology, together with the property models, the knowledge base, the databases and the design algorithm has to be highlighted through a number of case studies, involving different emulsified formulated products.

Figure 2.2 illustrates the objective of this PhD work, and the interactions between the integrated methodology and the developed methods and tools.
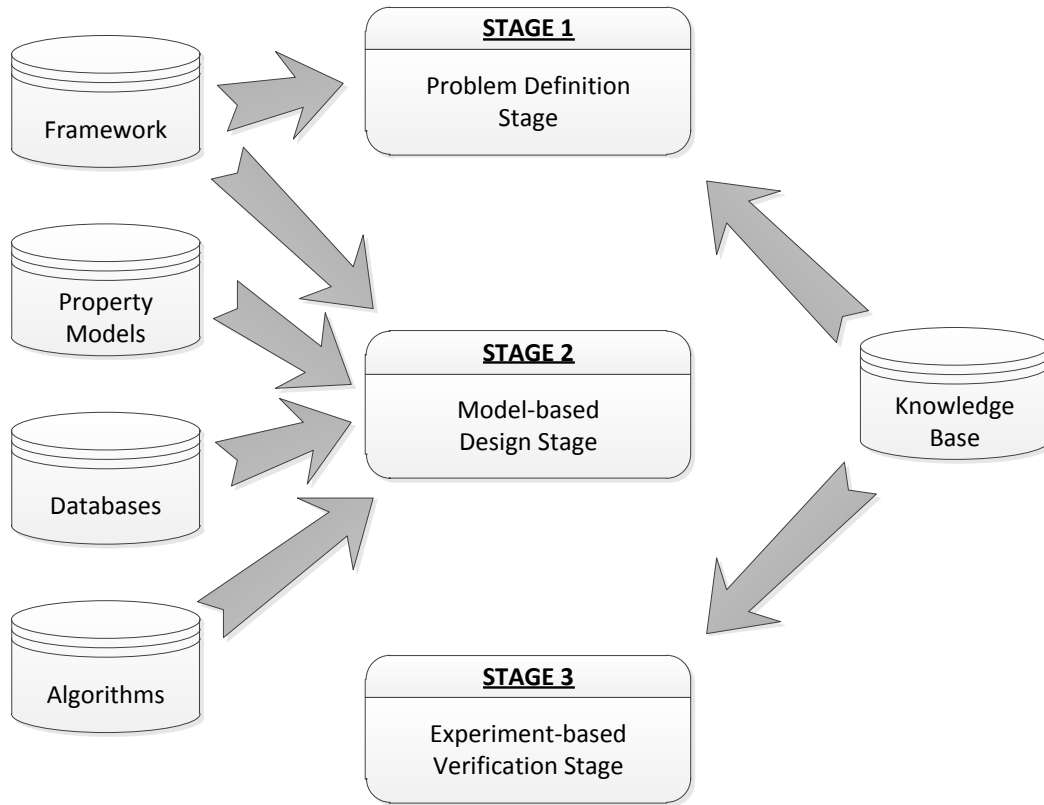
Figure 2.2 Summary of the objective of this PhD work and their interaction/integration with the methodology for emulsified formulated product design

3

PROPERTY MODELS

The objective of chemical product design, according to Gani (2004) is to find molecules, or mixtures of molecules with a desired set of target properties. Examples of single chemical, that is, molecular type, liquid products are solvents and process fluids, while mixtures of different chemicals are typically know as liquid formulated products. In the latter type of products, the chemicals responsible for providing the main function of the product (the active ingredients) are usually a solid in the range of conditions the product would be applied. Therefore, additional chemicals (the solvent mixture) are needed to dissolve and deliver the active ingredients. Other chemicals are then usually added to the formulation (the additives) to provide further enhancements of the product qualities.

The desired set of target properties represents the needs of the product, on the basis of which the design procedure is applied. Therefore, the product properties play a fundamental role in the design and development of chemical formulated products. It is through thermo-physical properties, in fact, that the stability of the product, the evaporation of the solvent mixture on application of the product, the spread-ability of the product, *etc.* are verified.

A common approach to design chemical products is by measuring the important properties for each candidate products. While this approach is reliable, it is also time consuming, expensive, and it does not allow the consideration of all the potential product candidates. A hybrid approach is usually recommended, where model-based techniques are used to estimate the desired set of properties and a set of promising candidates are identified through a model-based stage, while in a second experiment-based stage, the properties are verified and the formulated product is further improved, if necessary. In both cases,

however, a database of collected experimental data, supported by reliable mathematical models for prediction of thermo-physical properties is of fundamental importance.

In this chapter, an overview on the models for estimation of key thermo-physical properties for the design of emulsified formulated products is given. First, in §3.1, the properties are classified according to a modeling point of view; then, in §3.2, the property models are also classified, both from a property point of view, and from a mathematical perspective. A collection of the available property models is then presented according to the property type and the model type: first those that can be applied to a general liquid mixture (§3.3), followed by those property models that consider the presence of surfactants (§3.4). Finally, in §3.5, the property models developed in this work are presented.

It should be noted, however, that the property models considered or highlighted in this chapter are only examples and they are not necessarily the best and/or the most accurate for the corresponding thermo-physical property.

3.1 Classification of Properties

The selection of the most appropriate chemicals as active ingredients, solvent mixture and additives, for different types of formulated products, requires the evaluation of a wide range of thermo-physical properties. In fact, most of the product functions can be related to a set of thermo-physical properties. Table 3.1 gives a list of such relations.

Table 3.1 Product functions – target properties relationships

(In italic, those properties that are characteristics of emulsified formulated products)

Product Functions	Target Properties	Symbols
Conductivity	Dielectric constant	ε
Cost	Cost	C
Drying time	Evaporation time	T_{90}
Flammability	Open cup flash point	T_f
Foam-ability	*Critical micelle concentration, surface tension*	CMC, σ
Solubility	Hansen and Hildebrand solubility parameters	$\delta_D, \delta_H, \delta_P, \delta_T$
Skin irritability	Hansen and Hildebrand solubility parameters, pH	$\delta_D, \delta_H, \delta_P, \delta_T$, pH
Spray-ability	Density, dynamic viscosity, surface tension	ρ, μ, σ
Spread-ability	Density, kinematic viscosity	ρ, ν
Stability	Gibbs energy change of mixing	ΔG_{mix}
	Critical micelle conc., cloud point, Krafft temperature	*CMC, CP, T_k*
Toxicity	Toxicity parameter	LC_{50}

The set of desired target properties representing the product functions may be classified in terms of:

- Primary properties – These are single value properties of the pure compound. Every molecule is characterized by a single value of these properties. Examples are the critical properties, the normal boiling point, the normal melting point and many more. Measured values of these properties can usually be found in databases of chemicals. From a modelling point of view, they are also classified as those dependent only on the molecular structure of the compound.

- Secondary properties – From a modelling point of view, these properties of pure compounds are dependent on the molecular structure as well as other properties of the compound. For example, the density or heat of vaporization at the normal boiling point may be calculated from knowledge of the critical properties and the normal boiling point. In some cases, like the enthalpy of vaporization at the boiling point, a secondary property may be converted to a primary property.

- Functional properties (pure compound or mixture) – These are properties that depend on temperature, pressure and/or mixture composition. Those related to the pure component depend only on temperature and/or pressure, while those that depend also on mixture composition may be further classified as bulk-properties or compound properties in mixtures.

  Functional pure compound properties – These are properties such as vapour pressure, density, heat of vaporization, etc., of the pure compound that depend on the temperature and/or pressure. From a modelling point of view, the pressure effect is usually neglected and the temperature effect is modelled through regressed correlations. Details of the regressed correlations can be found for properties and compounds in databases of chemicals.

  Functional bulk properties – These properties are functions of mixture compositions as well as temperature and/or pressure (or a defined mixture state). They represent the bulk property of the mixture, for example, the density or viscosity of the liquid mixture (or formulation). From a modelling point of view, estimation of these properties require the corresponding pure compound properties of the involved compounds and a mixing rule to take into account the composition effect.

  Functional compound properties in mixtures – These are phase equilibrium related properties of the compounds present in a mixture. For example, the activity or fugacity coefficients of each compound present in the mixture. From a modelling point of view, they may require primary, secondary as well as other functional properties.

- Performance related properties – these properties are related to the performance of the product – such as the evaporation rate of the solvent and the stability of the liquid or emulsion. From a modelling point of view, they may require the above property models embedded into a process model. For example, the use of activity

coefficients (functional compound properties in a mixture) within a liquid phase stability test algorithm.

3.2 Classification of Property Models

The property models used to calculate the property values when not available in the literature may also be classified. For each class of properties, for example, it is possible to distinguish between those that are predictive by nature and those that are not. For example, estimating properties only from molecular structural information involve predictive models, such as the group contribution (GC) based models, while estimating properties from compound specific coefficients involve the use of correlations that are not predictive by nature. In liquid formulated product design, both types of models are needed.

Table 3.2 Target pure component properties, symbols
and property models applied in this work
(In italic, those properties that are characteristics of emulsified formulated products)

Target Properties	Symbols	Property Models
Cloud Point	CP	Group contribution method
Cost	C	Correlation
Critical micelle concentration	CMC	Group contribution method
Density	ρ	Group contribution method
Dielectric constant	ε	Correlation
Dynamic viscosity	μ	Group contribution method
Evaporation time	T_{90}	Correlation
Hansen solubility parameters	$\delta_D, \delta_H, \delta_P$	Group contribution method
Heat of vaporization	ΔH^{vap}	Correlation
Hildebrand solubility parameters	δ_T	Group contribution method
Hydrophilic-lipophilic balance	HLB	Definition
Kinematic viscosity	v	Definition
Krafft temperature	T_K	QSPR method
Open cup flash point	T_f	Group contribution method
Surface tension	σ	Group contribution method
Toxicity parameter	LC_{50}	Group contribution method
Vapor pressure	P^{vap}	Correlation

During the evaluation of candidate products, the models need to be predictive and computationally fast and cheap, while, during the verification of a small number of candidates, correlated models may be used, if the correlation coefficients are available. During the evaluation stage, the models need to be, at least, qualitatively correct, while, during the verification stage, the models also need to be quantitatively correct.

The property models may also be classified in terms of mechanistic (for example, applying quantum mechanical techniques) or correlative (for example, data based regression of correlations) or hybrid (for example, GC-based models that are predictive but require the use of regressed parameters). Truly predictive models are the mechanistic models but they are usually computationally expensive and as yet, their application range is limited. A vast variety of hybrid models have, however, been developed, ranging from simple easy to use GC-based models (Marrero, Gani, 2001; Fredenslund et al., 1977) to more complex "theoretical" models (Slater, 1951; Karayannis, Mavrantzas and Theodorou, 2004; Van Speybroeck, Gani and Meier, 2010). Other models that are not predictive with respect to the molecules that are present in the mixture but can extrapolate in terms of temperature, pressure and composition, are the well-known equations of state (Soave, 1972; Kontogeorgis et al., 1996) and the activity coefficient models (Fredenslund et al., 1977; Renon and Prausnitz, 1968; Klamt, 1995).

Table 3.3 Target properties, symbols and property models
applied in this work for homogeneous liquid mixtures

Target Properties	Symbols	Property Models
Cost	C	Linear mixing rule
Gibbs energy change of mixing	ΔG_{mix}	Activity coefficient models
Density	ρ	Linear mixing rule (on the molar volume)
Dielectric constant	ε	Linear mixing rule
Dynamic viscosity	μ	Linear mixing rule GC(UNIFAC)-based method
Hansen solubility parameters	$\delta_D, \delta_H, \delta_P$	Linear mixing rule
Hildebrand solubility parameters	δ_T	Linear mixing rule
Open cup flash point	T_f	Non-linear mixing rule
Surface tension	σ	Linear mixing rule GC(UNIFAC)-based method
Toxicity parameter	LC_{50}	Linear mixing rule
Vapor pressure	P^{vap}	Linear mixing rule Activity coefficient models

Table 3.2 provides a list of pure component properties, both primary and secondary, typically encountered in liquid formulated product design, together with the property models that can be applied. Tables 3.3 and 3.4, instead, list a collection of mixture properties, both functional bulk and compound in mixture, for the design of homogeneous and emulsified formulated products, respectively, with indication on the mixture models available. The properties and models illustrated in the above mentioned tables are described in the next sections, in §3.3 in relation to the common property models, and in §3.4 for those properties peculiar of the presence of surfactants.

Table 3.4 Target properties, symbols and property models
applied in this work for emulsified liquid mixtures

Target Properties	Symbols	Property Models
Cost	C	Linear mixing rule
Density	ρ	Linear mixing rule (on the molar volume)
Dielectric constant	ε	Two distinct values
Dynamic viscosity	μ	Correlation
Evaporation time	T_{90}	Two distinct values
Hansen solubility parameters	$\delta_D, \delta_H, \delta_P$	Two distinct values
Hildebrand solubility parameters	δ_T	Two distinct values
Hydrophilic-lipophilic deviation	HLD	Correlation
Open cup flash point	T_f	Two distinct values
Surface tension	σ	QSPR model
Toxicity parameter	LC_{50}	Linear mixing rule

3.3 General Property Models

In this paragraph, the first property models that have been used in this work are presented. In particular, those model that can be applied to any chemical or mixture of chemicals, independently by its nature are illustrated, under the definition of common property models. Within this paragraph, first the pure component property models are discussed in §3.3.1, followed by the mixture property models in §3.3.2.

3.3.1 Pure Component Property Models

The pure component property models have a dual values, when integrated in a methodology for formulated product design. On one hand, they are used to screen, evaluate and rank the candidate chemicals as active ingredients and additives, but on the other hand,

the pure component properties are also usually necessary for the estimation of the overall product properties by means of mixture property models.

Cost

The cost is not a thermo-physical property, nor strictly a property in general. However, it is a fundamental criteria for selecting the ingredients as well as the overall composition of a formulated products. The pure component cost data may be subject to a variety of uncertainties such as purity and source; therefore, Conte et al. (2011) proposed a simple correlation in order to provide qualitatively correct estimations of the pure compound cost, as a function of the molar volume, as in equations 3.1 and 3.2, with regards to alcohols and esters, respectively.

$$C = 2.152 \cdot V_m - 38.714 \tag{3.1}$$

$$C = 2.356 \cdot V_m - 119.00 \tag{3.2}$$

Where C is the cost in \$/kmol, and V_m is the molar volume expressed in L/kmol.

It is necessary to notice, however, that this model is not accurate, as it does not take into account the fluctuations of the market, and it has been developed only for preliminary selection purposes, when the cost of several potential candidates cannot be retrieved.

Density

The liquid density is a fundamental thermo-physical property in the design of liquid formulated products. In fact, since it relates not only to the volume per mass of the product as well as to several secondary and functional properties, but it is also a key parameter for the sizing of the processing equipment.

Most of the prediction methods for saturated liquid densities are based on the corresponding state principle. One of the most popular methods for the prediction of liquid densities is the Rackett equation (Rackett, 1970), further modified by Spencer and Danner (1972) and by Yamada and Gunn (1973). The modified version of the Rackett equation proposed by Spencer and Danner is the most commonly used model, and it is given in equation 3.3.

$$\frac{1}{\rho_L} = \left(\frac{R \cdot T_{crit}}{\rho_{crit}} \right) \cdot Z_{Ra}^{\left[1 + (1 - T_R)^{2/7} \right]} \tag{3.3}$$

Where ρ_L is the liquid density, T_{crit} is the critical temperature, ρ_{crit} is the critical density, T_R is the reduced temperature, R the universal gas constant and Z_{Ra} is the constant of the modified Rackett equation. This needs to be determined from experimental data, either the critical compressibility factor (Z_{crit}), given in equation 3.4 can be used instead.

$$Z_{Ra} = Z_{Crit} = \frac{P_{Crit} \cdot V_{Crit}}{R \cdot T_{Crit}} \tag{3.4}$$

Where, P_{crit} is the critical pressure and V_{crit} is the critical volume.

In addition, group contribution based models, which only need the molecular structural information of the chemical, have been developed for the prediction of pure compound liquid densities, through the molar volumes, at the standard state. The GCVOL model by Elbro, Fredenslund and Rasmussen (1991) and its extensions by Tsibanogiannis, Kalospiro and Tassios (1994) and by Imhels and Gmehling (2003) are good examples of this type of models.

In general, however, the liquid density is calculated as the inverse of the liquid molar volume: a state variable that is found in all the equations of state. That is, from any equation of state, given two state variables, such as, the temperature and the pressure, the corresponding liquid molar volume, and therefore the liquid density can be calculated.

Dielectric Constant

The dielectric constant is an important property in the formulated product design, as this property is correlated to a number of secondary thermo-physical properties, such as, the solubility and the polarizability.

The most commonly applied method for the calculation of this property is a correlation reported by Horvath (1992), which is suitable for a broad variety of chemicals. The mathematical expression is given in equation 3.5.

$$\varepsilon = \frac{\delta_T - 7.5}{0.22} \tag{3.5}$$

Where ε is the dielectric constant, and δ_T is the Hildebrand solubility parameter.

However, it is possible to calculate the dielectric constant of a pure component through group contribution based models, modelling this property, therefore, as a primary property (Sheldon, Adjiman and Cordiner, 2005; Megnassan, Legoff and Proutiere, 1994; Gani, Harper and Hostrup, 2005). In this type of method, extensively described in the Appendix A, with respect to the one developed by Marrero and Gani (2001), a chemical is uniquely represented by a set of functional groups. Once the group contributions for a corresponding property are retrieved from the group contribution table, the property is estimated through an expression like equation 3.6, representative of the method by Constantinou and Gani (1994).

$$F(\zeta) = \sum_i N_i C_i + \sum_j M_j D_j \tag{3.6}$$

Where C_i is the contribution of the first-order group of type i, which occurs N_i times; and D_j is the contribution of the second-order group of type j, which occurs M_j times.

Dynamic and Kinematic Viscosity

The knowledge of the dynamic viscosity of a pure component plays an important role in those design issues related to transport of mass and/or energy, and many more properties such as the spread-ability and the spray-ability of the liquid product, as from table 3.1.

Among the several models proposed, the one proposed by Sastri and Rao (1992) is commonly applied, as this correlation is reliable for different types of organic chemicals. The mathematical expression is given in equation 3.7.

$$\mu = \mu_B \cdot P^N \qquad (3.7)$$

Where μ is the dynamic viscosity, expressed mPa/s, P is the room pressure in atm, and μ_B and N are regressed parameters that are characteristics of each type of chemicals, as they are found to vary very little for compounds with a similar molecular structure, therefore Sastri and Rao propose a group-contribution method for the estimation of such parameters. Recently, a group-contribution method has been developed for the prediction of the dynamic viscosity at 298 K by Conte, et al. (2008), and it has been found to give better performances.

The kinematic viscosity, on the other hand, is calculated on the basis of its definition, as the ratio between the dynamic viscosity and the liquid density, as in equation 3.8.

$$\nu = \frac{\mu}{\rho} \qquad (3.8)$$

Environmental and Health Related Properties

There has been raising interest, in the last decades, on the prediction of environmental and health related properties. Such properties, in fact, are crucial in the product development, as they influence the possibility for a product to be sent in the market.

Among different models, a series of group-contributions models have been recently developed by Hukkerikar et al. (2012) for a range of environmental, health and safety related properties. The following properties were covered: the fathead minnow 96-hr LC_{50}, daphnia magna 48-hr LC_{50}, oral rat LD_{50}, aqueous solubility, bio-concentration factor, permissible exposure limit (OSHA-TWA), photochemical oxidation potential, global warming potential, ozone depletion potential, acidification potential, emission to urban air (carcinogenic and non-carcinogenic), emission to continental rural air (carcinogenic and non-carcinogenic), emission to continental fresh water (carcinogenic and non-carcinogenic), emission to continental sea water (carcinogenic and non-carcinogenic), emission to continental natural soil (carcinogenic and non-carcinogenic), emission to continental agricultural soil (carcinogenic and non-carcinogenic).

In this paragraph, the LC_{50} property is highlighted, which is among the most used target properties in process and product design. The parameter LC_{50} indicates the lethal concentration of a pure chemical or mixture that causes 50% of deaths in a fathead minnow population. Especially in the design of formulated products, such as a skin-care or a cosmetic product, or products to be inhaled or ingested, this parameter is used as a measure of the toxicity along with other properties listed above.

The model parameters for all the listed properties for the M&G GC+ method (see Eq. 6) are given by Hukkeriker et al. (2012). A first-order group-contribution model is used, as given in equation 3.9.

$$-\log(LC_{50}) = \sum_i N_i C_i \qquad (3.9)$$

Evaporation Time

The evaporation time is needed in the design of liquid formulated products when the solvent is supposed to vaporize out after application of the product. Experimental values for the evaporation time are difficult to measure and poorly available in the literature; therefore, correlations based on model systems reported by others are usually employed.

Van Wesenbeck, Driver and Ross (2008) proposed a linear correlation between the evaporation rate (ER) and the vapour pressure (P^{vap}), as in equation 3.10.

$$\ln(ER) = 0.865 \cdot \ln(P^{vap}) + 12.70 \qquad (3.10)$$

Where the evaporation rate is expressed in g·cm^2/s and the vapour pressure in Pa.

Klein, Wu and Gani (1992) found the above mentioned correlation to estimate the evaporation rate of a pure solvent quite accurately, but other correlations as well have been proposed (Conte, 2010) and they are reported to give a good qualitative accuracy in the description of the relation between the vapour pressure of the solvent and its evaporation rate.

Very often, instead of the evaporation rate, the target property is the evaporation time (T^{90}, expressed in s), which is the time needed for 90% by weight of the solvent to evaporate. This property is also correlated as a function of the vapour pressure of the solvent, as from equation 3.11.

$$\ln(T^{90}) = -0.793 \cdot ln(P^{vap}) + 12.416 \qquad (3.11)$$

Conte, Gani and Ng (2011) found also the accuracy of this correlation to be quite satisfactory.

Heat of Vaporization and Vapour Pressure

The heat of vaporization and the vapour pressure of pure components are classified as functional pure compound properties, dependent on temperature. This means that the measured data of these properties are correlated for each compound as functions of temperature. The regressed coefficients can be found in databases of chemicals and their properties (Nielsen et al., 2001; Kroenlein et al., 2011; DIPPR, 2003). Also, equations of state, such as the cubic equations of state (Soave, 1972) could be used if the necessary critical properties of the compounds are known. For non-cubic equations of state, the parameters, for example, of CPA (Kontogeorgis et al., 1996) and PC-SAFT (Gross and Sadowski, 2001) are estimated on the bases on vapour pressure data. Another option requiring only the molecular structural information, and that is predictive in nature, is represented by group contribution models, to estimate the regressed parameters of the temperature dependent functions (Ceriani, Gani and Meirelles, 2010).

The modeling of the vapor pressure and of the heat of vaporization of lipids is finding increasing use in chemicals based liquid products, and Ceriani, Gani and Meirelles (2009)

developed a group contribution model using an extensive database of lipids of various classes with improved predictive power.

The property model for the vapor pressure and for the heat of vaporization is given by equations 3.12 and 3.13, respectively.

$$\ln(P^{vap}) = A + \frac{B}{T^{1.5}} - C \cdot \ln(T) - D \cdot T \tag{3.12}$$

$$\Delta H^{vap} = -R \cdot \left(\frac{1.5 \cdot B}{\sqrt{T}} + C \cdot T + D \cdot T^2\right) \cdot \left(1 - \frac{T_{Crit}^3 \cdot P^{vap}}{T^3 \cdot P_{Crit}}\right) \tag{3.13}$$

Where P^{vap} is the vapor pressure in Pa, T is the temperature in Kelvin, ΔH^{vap} is the heat of vaporization in J/mol, T_{Crit} and P_{Crit} are the critical temperature and pressure, respectively, and A, B, C and D are model parameters whose values, obtained from the regression of experimental data. Are given by Ceriani, Gani and Meirelles (2009).

Open Cup Flash Point

The open cup flash point is related to the flammability of a chemical and therefore is included as a target property to accommodate safety issues. Since the solvent mixture is supposed to evaporate after application of the product, the flammability issue is as important as environmental and health impacts. The solvent mixture should have a flash point which is at least higher than the usage temperature of the product, considering that in the formulation the solvent mixture is diluted by active ingredients and additives that are usually not highly flammable.

Here also, group-contribution models have been developed in order to predict this important property with acceptable accuracy, by Constantinou and Gani (1999). Moreover, quantitative structure-property relationship (hence, QSPR) models have also been developed, the most comprehensive of which has been proposed by Gharagheizi, Keshavarz and Sattari and its mathematical expression is illustrated in equation 3.14.

$$T_f = 43.512 + 0.8374 \cdot NBP + 1.635 \cdot S_s - 39.1658 \cdot vEv1 \tag{3.14}$$

Where T_f is the flash point in K, NBP is the normal boiling point (also, in Kelvin), while S_S and vEv1 are molecular descriptors of the model.

Solubility Parameters

The solvent selection is one of the major concerns in the early development of many formulated products, because of the relation with the active ingredients and the additives in the product. Since the active ingredients and additives, in their standard states, are usually in the solid form, solvents are needed to dissolve them and then to deliver them to the product application site.

The prediction of the solid solubility is a major topic of current research and many method and tools are available for the estimation of solubility of solids in solvents (Karunanithi, Achenie and Gani, 2007). The solid-solvent systems may be quite complex (Ablidskov and O'Connel, 2003) and the applicability of any single method with acceptable accuracy

for a wide range of chemical systems is questionable (Gani, Jimenez-Gonzalez and Constable, 2005).

A model-based liquid formulated product design technique would require the screening of thousands of candidate solid-solvent mixtures; therefore, a predictive method, that is also easy to apply, is needed. A good review on model-based estimation of solid solubility is given by Conte et al. (2008). In this paragraph, however, the use of the solubility parameters for the solvent selection is discussed.

At the level of early stage design, in fact, it is useful to identify the best solvents without detailed calculations of the solid-solvent solubility. For this purpose, the Hildebrand solubility parameter (δ_T) has been introduced by Hildebrand (1936), for fast, efficient screening of solvents. It is defined as in equation 3.15.

$$\delta_T = \left(\frac{\Delta H^{vap}}{V_m}\right)^{0.5} \tag{3.15}$$

The reliability of the decision taken on the basis of the Hildebrand solubility parameters, however, are questionable as it does not take into account the variations in the contributions to the heat of vaporization due to different effect, such as, dispersion, polar and hydrogen-bonding forces. Therefore, the Hansen solubility parameters may be used instead. These, defined by Hansen (1967) as δ_D as the dispersion parameter, δ_P as the polar parameter, and δ_D as the hydrogen bonding parameter, give a qualitative understanding of the solvent issues, thereby allowing a fast screening between suitable and unsuitable candidate solvents, significantly reducing the search space and the computational load for more rigorous quantitative approaches. A group-contribution predictive model based on the work of Marrero and Gani (2001), the mathematical expression of which is reported in equation 3.16, has been proposed by Modarresi et al. (2008).

$$F(\zeta) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k \tag{3.16}$$

Where, compared to equation 3.6, E_k is the contribution of the third-order group of type k, which occurs O_k times.

Later, Hukkerikar et al. (2012) confirmed that through this model, each of the three Hansen solubility parameters can be predicted with good accuracy. Clearly, there is a straight relation between the Hildebrand and the Hansen solubility parameters, as described in equation 3.17.

$$\delta_T = (\delta_D^2 + \delta_P^2 + \delta_H^2)^{0.5} \tag{3.17}$$

Surface Tension

Like the dynamic viscosity, the surface tension is also a key property in process and product design since it strongly influences several transport properties and performances, such as, the spray-ability among the others. In the past, the method developed by Orrick and

Erbar (personal communication) has been used, and also QSPR models have been proposed (Delgado and Diaz, 2006) providing good results with reasonable accuracy for a wide range of chemicals. However, these days, the use of a group-contribution plus connectivity index based method, described in equation 3.18, as proposed by Conte et al. (2008), is an alternative worth considering for prediction of surface tension of a very wide range of chemicals, at constant temperature.

$$F(\zeta) = \sum_i N_i C_i + F(\zeta^*) + \sum_j M_j D_j + \sum_k O_k E_k \qquad (3.18)$$

Where, differently than equation 3.16, $F(\zeta^*)$ is a function of the surface tension for all the groups, the contribution of which is not available in the parameter database. It represents the part of the connectivity index based method.

3.3.2 Linear Mixture Property Models

As mentioned at the beginning of this chapter, most of the product target properties are relative to the formulated liquid mixture, often referred to as bulk properties of the mixture. The mixture property models, from a mathematical perspective, can be classified into linear and non-linear.

The simplest model for the estimation of the bulk mixture properties is to use the corresponding pure compound properties and a linear mixing rule. The question of when to apply the linear mixing rules depends on the type of the chemicals present in the mixture. For mixtures with negligible excess properties of mixing the linear mixing rules may be safely used. For others, depending on the accuracy needed and the availability of other models, the linear mixing rules may also be employed, as an initial trial.

According to the linear mixing rule, the bulk mixture property is estimated through equation 3.19, where the mixture property ζ, is determined from the pure compound properties (ζ_i) and mole fractions (x_i) of the compounds present in the mixture, at a specific temperature and pressure.

$$\zeta = \sum_i \zeta_i x_i \qquad (3.19)$$

A good indication of mixtures that may show negligible excess properties of mixing can be obtained from plots of excess properties of known binary mixtures, as given by Smith, Van Ness and Abbott (2001).

Note that for some properties, such as the liquid density, even though linear mixing rule cannot be strictly applied, the molar volume has been observed to behave linearly with composition. Thus, it is possible to estimate the molar volume at a fixed temperature and pressure, through the linear mixing rule and then convert it to the density. A list of the target properties for which the linear mixing rule is usually applied is given in table 3.3.

3.3.3 Non-linear Mixture Property Models

For some bulk properties of the mixtures, instead, the calculation methods employing the linear mixing rule cannot be used. In this paragraph, these properties and the relative non-linear mixture property models are given.

Dynamic Viscosity

For a given phase of a mixture together with its temperature and pressure, non-linear mixing models that are also based on group contributions have been developed for the estimation of bulk mixture properties, such as the dynamic viscosity. Cao et al. (1993) developed, in fact, a very useful predictive model, based on group-contribution models, for the calculation of the bulk mixture dynamic viscosity, once the pure component values are available. The model is described in equation 3.20.

$$\ln(\mu \cdot V_m) = \sum \varphi_i \cdot \ln(\mu_i \cdot V_{m,i}) + 2 \sum \varphi_i \cdot \ln\left(\frac{x_i}{\varphi_i}\right) + \tag{3.20}$$
$$- \sum \frac{n \cdot q_i \cdot p_i \cdot \varphi_i}{r_i} \cdot \sum \theta_{ij} \cdot \ln(\tau_{ij})$$

Where p_i, q_i, r_i, n, φ_i, θ_{ij} and τ_{ij} are parameters of the UNIFAC models and therefore, as long as these group parameters are available in the UNIFAC table, this model can be used.

Open Cup Flash Point

The estimations of the open-cup flash point, as a measure of the flammability of a mixture, requires models suitable for non-ideal mixtures depending on the chemicals present in the mixture. A reliable mathematical model for predicting this property has been developed for miscible mixtures by Affens and McLaren (1972), and it is given in equations 3.21 and 3.22.

$$1 = \sum \frac{10^a \cdot x_i \cdot \left[1642 - \left(T_{f,i} + 230\right)\right]}{\left[1642 - \left(T_{f,mix} + 230\right)\right]} \tag{3.21}$$

$$a = \frac{m_i\left(T_{f,i} - T_{f,mix}\right)}{T_{f,i} \cdot T_{f,mix}} \tag{3.22}$$

Where $T_{f,i}$ and $T_{f,mix}$ are the pure component and the mixture flash point, respectively, while m_i is a model parameter. The results, provided that accurate flash point values of the pure compound thermo-physical properties are available, confirm that this model can be reliably applied for process-product design issues.

Surface Tension

Similarly to the mixture model relative to the dynamic viscosity, also for the surface tension, a predictive model, based on the UNIFAC activity coefficient model, has been developed and presented by Suarez, Torres-Marchal and Rasmussen (1989) and its mathematical expression is give in equation 3.23.

$$\sigma_{mix} = \sigma_i + \frac{R \cdot T}{A_i} \cdot \frac{x_{i,s} \cdot \gamma_{i,s}}{x_{i,b} \cdot \gamma_{i,b}} \qquad (3.23)$$

Where γ is the activity coefficient, σ_i and σ_{mix} are the pure component and the mixture surface tension, and the sub-script s and b stand for the surface and bulk liquid phase, respectively. In fact, as most of the models for the description of the surface tension, this model is based on the assumption that the surface can be treated as a separate phase located between the vapor and the bulk liquid phase. Finally, A_i is a model parameter that can be predicted if all the necessary UNIFAC parameters are available.

The accuracy of this model is mostly acceptable, but for highly non-ideal system, such as, electrolyte systems, the errors may be large.

Phase Equilibria-related Properties

Many of the thermo-physical properties for which a linear mixing rule cannot be applied, can be calculated from the corresponding conditions of the phase equilibrium. Examples are the above mentioned dynamic viscosity and the surface tension, but also the vapor pressure (from which the heat of vaporization and the evaporation time are calculated) and the Gibbs energy change of mixing. Finally, also the solubility can be quantitatively described by means of activity coefficient models and/or equations of state.

The starting points are the conditions of phase equilibrium: the vapor-liquid equilibrium (equation 3.24), the liquid-liquid equilibrium (equation 3.25) and the solid-liquid equilibrium (equation 3.26).

$$y_i \varphi_i^V P = x_i \gamma_i^L P_i^{vap} \qquad (3.24)$$

$$x_i^I \gamma_i^{L,I} = x_i^{II} \gamma_i^{L,II} \qquad (3.25)$$

$$exp\left[\frac{\Delta H_i^{fus}}{R \cdot T_m} \cdot \frac{T - T_m}{T}\right] = x_i \gamma_i^L \qquad (3.26)$$

Where φ_i is the fugacity coefficient, ΔH^{fus} is the heat of fusion and T_m is the melting temperature. The super-scripts I and II of equation 3.25, moreover, indicates that the properties refer to one or the other liquid phase.

At a given composition and pressure, the above equations may be used to calculate the temperatures at which the conditions of equilibrium are satisfied. The pure compound vapour pressures and the normal melting points and enthalpy of fusions are obtained from the models discussed in §3.3.1. The estimations of the needed fugacity and the activity coefficients, instead, introduces the concept of activity coefficient models and equations of state.

The most well-known models for the estimation of the fugacity coefficient are the cubic equations of state (Soave, 1972; Van der Waals, 1873; Peng and Robinson, 1976), which are, however, not particularly accurate for very non-ideal mixtures. Evolutions of the above mentioned equations of state are the CPA (Kontogeorgis et al., 1996) and the PC-SAFT (Gross and Sadowski, 2001), are able to perform well for a wide range of mixtures.

Regarding the activity coefficient, instead, the most commonly used models are those based on the molecular interactions, such as, those proposed by Wilson (1964), Renon and Prausnitz (1968), and Abram and Prausnitz (1975). The UNIFAC model, proposed by Fredenslund et al. (1977), is predictive, as long as the necessary group parameters are available. For this reason, it is often used at the early stage of process and product design, when it is not worth to collect all the experimental data that are necessary to apply the other models.

3.4 Surfactant-specific Property Models

The surfactants are chemicals characterized by an amphiphilic nature, that is, part of the molecule is hydrophilic, while another part is hydrophobic. In order to describe their behaviour in emulsified formulated products, some very peculiar properties, such as, the cloud point, the critical micelle concentration, *etc.* are needed. These properties are specific of this type of molecules, therefore they are not described in §3.3.

However, the knowledge of surfactants and emulsions is not as developed as the homogeneous case, and therefore predictive models for many of the needed properties are often lacking. Part of this work, then, consisted of the development of group-contribution based methods, for the prediction of surfactant-specific property models. In this paragraph, existing as well as newly developed models are presented together, while details on the in-house models developed and proposed in this work are given in §3.5.

The choice of the cloud point and the critical micelle concentration as target properties to be modelled with group-contribution methods is due to their importance in the choice of the candidate surfactants to be included in the emulsified solvent mixture, as well as in the definition of the minimum necessary concentration of it. Moreover, the proposition of a thermodynamic-based method to assess the stability of an emulsion, as a function of composition and temperature, comes from the high uncertainty that characterizes the correlation widely used for the same purpose.

3.4.1 Pure Component Property Models

Although many of the surfactant-specific properties cannot strictly be considered as pure component properties, they refer to mixtures where the temperature and/or the composition are fixed. Therefore, these properties can be obtained from the only information relative to the molecular structure of the surfactant, so that they are modelled as primary properties and they can be estimated using predictive models such as the QSPR and the group-contribution models.

Cloud Point

One of the surfactant-related properties that are considered fundamental for the development of a methodology for the design of emulsified formulated product, is the cloud point. This property, sometimes referred to as the cloud temperature, is specific of mixtures

between water and nonionic surfactants and Mitchell et al (1983) defined it as the temperature at which the mixture starts to phase separate and two phases appear, thus becoming cloudy. This phenomenon is of particular relevance for those non-ionic surfactants containing polyoxyethylene chains, exhibiting reverse solubility versus temperature behaviour. This phenomenon is illustrated in Figure 3.1, where a standard phase envelope of an aqueous mixture of a polyoxyethylene-based nonionic surfactant is given.
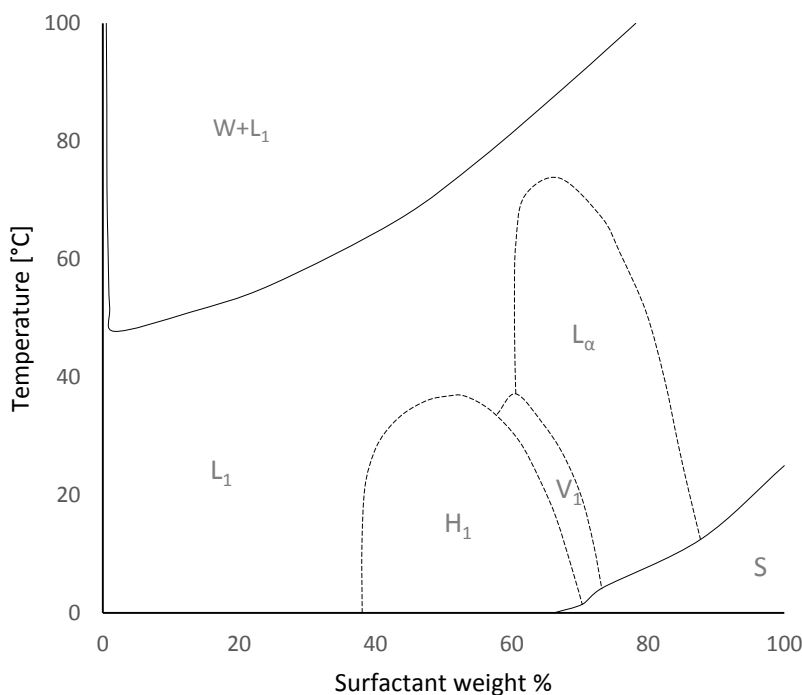
Figure 3.1 Phase envelope of an aqueous mixture of a polyoxyethylene-based nonionic surfactants. Data are taken from Mitchell et al. (1983).

In Figure 3.1, different regions can be recognized: L_1 identifies an aqueous surfactant solution where the surfactant is organized in ordinary or reverse spherical micelles; W represents a very diluted surfactant solution (around the critical micelle concentration); S indicates the presence of solid surfactant; L_α, H_1 and V_1, instead, are regions where the surfactant is aggregated in unusual structures such as, respectively, lamellar, normal hexagonal and bi-continuous cubic structures. The last three regions are sometimes grouped together as "viscous" meso-phases, since their rheological properties and behaviour are substantially different from those of the ordinary and reverse spherical micellar solution. The line separating the $W+L_1$ area from the L_1 area is the cloud point line.

It is evident that the cloud point does not depend only on the system considered, but it is influenced by the surfactant content. However, it is common practice (Guo et al., 2012) to define as the cloud point the numerical value assumed by the cloud point curve at a surfactant weight percentage of 1%.

QSPR models have been extensively applied for the modelling of the cloud point (Ren et al., 2011), but group-contribution models are usually preferred in chemical process and

product design as they provide accurate predictions without being computationally demanding. Moreover, they can be easily integrated in a computer-aided molecular design technique, as they employ the same building block for the molecular representation.

In this work, the group-contribution method presented by Marrero and Gani (2001) has been applied. A total data-set consisting of 86 experimental cloud points of different surfactants has been collected, involving several families of non-ionic surfactants. A thorough data analysis has recommended a particular choice of the form of the f(ζ) of equation 3.16, and it is illustrated in equation 3.27.

$$CP^2 = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k \qquad (3.27)$$

Newly defined third order groups have been introduced, thanks to a systematic data-error analysis, for the description of the data-set.

The experimental data, the calculated data, and the model parameters are reported in Appendix A, while more details about the model development are given in §3.5.2.

The availability of reliable experimental data regarding the cloud points of other families of non-ionic surfactants, such as, alkanediols, esters, ethers and fluorinated polyoxyethylene surfactants can broaden the application range of the model. However, it can already be safely applied for the prediction of the cloud point, considering the limited maximum error registered, with poor limitations on the molecular structures accepted.

Critical Micelle Concentration

The critical micelle concentration is defined as the limit below which virtually no micelles are detected, and the limit above which virtually all the additional surfactant molecules form micelles. Many properties of the surfactants, if plotted versus the concentration, appear to change at a different rate above and below this value; the critical micelle concentration is then considered a fundamental property of the surfactants. Ray (1971) demonstrated that only mixtures of surfactants and polar solvents with two or more potential hydrogen-bonding sites are capable of forming micelles, while in non-polar solvents, clusters of surfactants may form, but their behavior is not comparable with that of micelles in aqueous media.

Several contributions can be found in the literature, in relation to the modeling of the critical micelle concentration of non-ionic surfactants: from simple correlations (Klevens, 1953), to molecular simulation techniques (LeBard et al., 2012), to activity coefficient models (Chen, 1996; Voutsas et al., 2001), to equations of state (Li, Lu and Li, 2000). Results are in general encouraging, but experimental data are necessary for estimating the model parameters, and therefore the extrapolation of such models for predictions is not recommended. Among the predictive models, instead, several QSPR models have been developed. An exhaustive, extensive review has been recently given by Hu, Zhang and Wang (2010).

The group-contribution model developed and further applied in this work is inspired to the method proposed by Marrero and Gani (2001). The data-set used for the parameter regression consisted of 120 experimental data-points, 30 of which have been used for model validation purposes, representative of a very wide range of non-ionic surfactants: linear, branched and phenyl polyoxyethylene-based surfactants, alkanediols, alkyl mono and disaccharide ethers and esters, ethoxylated alkyl amines and amides, fluorinated linear ethoxylates and amides, poly-glycerol esters and carbohydrate derivate ethers, esters and thiols.

From the analysis of the experimental values of the critical micelle concentration, in relation to the molecular structures of the surfactants considered, the choice of the form of the f(ζ) of equation 3.16 is reported in equation 3.28.

$$-\log(CMC) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k \qquad (3.28)$$

Newly defined third-order groups have been then required for the description of such a wide range of molecular structure.

The experimental data, the calculated data and the model parameter are reported in Appendix B, while details on the model development and on the model performances are reported in §3.5.3. However, the performances of the model developed in this work are satisfactory, and some high prediction errors that are observed for some families of surfactants can be attribute to the insufficient availability of data for such groups of surfactants. As a future improvement of the method, the inclusion in the data-set of more experimental values is recommended, as the application range can be widened.

The application of a group-contribution method as above to the modelling of the critical micelle concentration of ionic surfactants is currently a work in progress. In fact, because of the presence of ionic groups in the molecular structures, first and second-order groups need to be redefined accordingly. The preliminary results, however, are found to be promising. As long as a reliable group-contribution model is not available, the QSPR model developed by Wang et al. (2002), given in equation 3.29, is considered the state of the art, for the prediction of the critical micelle concentration of ionic surfactants.

$$\log(CMC) = 0.546 - 0.269 \cdot KHO - 0.0037 \cdot \Delta H^{form} + 0.382 \cdot E_{HOMO} \qquad (3.29)$$
$$+ 0.493 \cdot E_{LUMO} - 0.0134 \cdot D$$

Where ΔH^{form} is the heat of formation, D the molecular dipole moment, and KHO, E_{HOMO} and E_{LUMO} are molecular descriptors of the model.

Hydrophilic-Lipophilic Balance

The hydrophilic-lipophilic balance (HLB) system is probably the most common method to correlate the surfactant molecular structures with their performances as emulsifiers. This system, first developed by Griffin (1949) consists of an arbitrary scale, from 1 to 20, to which HLB numbers are experimentally determined and assigned. The value of the HLB number indicates how the surfactant will behave in a solution and which kind of

emulsion would be formed by adding that surfactant to a two non-miscible phase system. If the HLB value is lower than 10, the surfactant is hydrophobic, therefore a water-in-oil emulsion can be formed; while if it is higher than 10, the surfactant is preferably soluble in water, and therefore an oil-in-water emulsion can be produced. The HLB concept works very well for non-ionic surfactants, but it is less successful with ionic surfactants. The most general model for the estimation of the HLB has been proposed by Davies (1957) and it is a group-contribution based model, illustrated in equation 3.30.

$$HLB = 7 + n_{i,h} \cdot C_{i,h} - n_{i,l} \cdot C_{i,l} \tag{3.30}$$

Where n is the number of groups of type i in the molecule, and C is the respective contribution. The sub-scripts h and l, instead, refer to the hydrophilic and lipophilic groups, respectively.

Alternative methods to describe the amphiphilic behaviour of surfactants have been proposed by Beerbower and Hill (1971) and by Israelachvili, Mitchell and Ninham (1976), but even if the HLB system may sometimes be found inaccurate, it is still preferred in the early stage process and product design for surfactant screening and selection.

Krafft Temperature

Surfactants forming micelles exhibit unusual solubility behaviour as their solubility shows a rapid increase above a certain temperature, known as the Krafft point or the Krafft temperature. This solubility behaviour is mostly observed for ionic surfactants as only a few non-ionic surfactants possess a real Krafft temperature (Salager, 1996). Duncan (1983) observed that the Krafft temperature depends on a complex three phase equilibrium condition and it largely depends on the counter-ion of the surfactant, on the length of the hydrophobic moiety, and on the presence of electrolytes.

These experimental evidences make the development of dedicated models for Krafft temperature complicated. So far, a correlation with limited application range has been proposed by Gu and Sjoblom (1992), but the state of the art is represented by the QSPR model proposed by Li et al (2007), reported in equation 3.31.

$$T_K = 57.4 - 7.6 \cdot KS2 - 0.06 \cdot \Delta H^{form} + 47.1 \cdot A \cdot \log(P98) - 28 \cdot A \tag{3.31}$$
$$\cdot \log(P) - 36.1 \cdot IC + 6.7 \cdot nO$$

Where log(P) is the octanol-water partition coefficient, while KS2, A, P98, IC and nO are the model descriptors.

Similarly to what mentioned in relation to the estimation of the critical micelle concentration of ionic surfactants, group-contribution based models for the estimation of the Krafft temperature are under development, and the preliminary results are encouraging.

3.4.2 Mixture Property Models

When considering emulsified formulated products, it is necessary to note that, even though from a macroscopic point of view the product may appear to be homogeneous, at the microscopic level emulsions consist of two non-miscible phases, kept together by the

presence of adequate surfactants. The bulk mixture property models presented in §3.3.2 and §3.3.3 may not be suitable for the estimation of the thermo-physical properties of emulsions. In this paragraph, the mixture property models in relation to emulsified formulated product are discussed.

First, those mixture properties that can be described with the linear mixing rule (§3.3.2) are highlighted, followed by those properties for which two distinct values characterize the emulsified formulated product. Finally, those properties requiring dedicated models that consider the specific nature of the emulsion are presented.

Linear Mixing Rule

The linear mixing rule is known to be suitable only when the excess mixture properties are not significant. When considering an emulsion, since they consist of two non-miscible phases, such an assumption is hardly verified. However, many target mixture properties can be estimated through this simple model, such as, the cost, the overall density, and the toxicity parameter. In the specific case of emulsified formulated products, the equation of the linear mixing rule can be slightly re-written, as in equation 3.32

$$\zeta = \zeta_A \cdot x_A + \zeta_O \cdot x_O \tag{3.32}$$

Where the sub-script A and O refer to the aqueous and to the organic solvent phase, respectively.

Two Distinct Values

Because of the amphiphilic nature of the emulsified products, where two non-miscible phases coexist in a mixture which is homogeneous from a macroscopic perspective, some mixture properties cannot be quantified by one single values. As a simple example, there is not such a property as the evaporation time of en emulsified product, since the organic solvent phase is characterized by its evaporation time, as well as the aqueous solvent phase has its own evaporation time. As a consequence, the two phases of an emulsified formulated product are subject to different evaporation rates and if the difference between the values of the two phases is large enough, it is possible that one of the two phases completely evaporate well before the other one, thus influencing the stability of the emulsion as well as many other thermo-physical properties. Also, considering the solubility parameters, an emulsified formulated product is described by two sets of solubility parameters, one per each solvent phase.

Among the properties listed in Table 3.4, the dielectric constant, the evaporation time and the solubility parameters belong to this category of mixture properties.

Non-linear Mixing Rules

There is a set of mixture properties, however, that needs dedicated models able to manage the emulsified form of the formulated product. These are the dynamic viscosity, the hydrophilic-lipophilic deviation, and the surface tension.

The dynamic viscosity of an emulsions, in fact, depends on a complex function of the properties of the two solvent phases and their relative fraction. Strictly, it is also influenced by the droplet size distribution, both in terms of average value and in terms of dispersion of the distribution, but such an effect is considered negligible for most of the commercial emulsified formulated products.

In the recent few years, several modelling efforts have been done in order to develop reliable models that are now widely used, especially in the field of the design and production of creams and pastes. However, due to characteristic high uncertainties, the model validation by means of experimental is usually required.

It has to be noticed that emulsions show in general a very non-ideal rheological behaviour, making the assumption of Newtonian fluid completely unrealistic. Depending on the formulation variables, emulsions can be defined both as pseudo-plastic and as pseudo-dilatant fluids. However, the models developed are set to evaluate the dynamic viscosity in term of an apparent viscosity, based on Newton's postulate.

For diluted emulsions, the model proposed by Bartok and Mason (1958) is usually applied, and it is reported in equation 3.33.

$$\mu = \mu_S \cdot \left[1 + \left(\frac{1 + 2.5 \cdot M}{1 + M}\right)\psi\right] \qquad (3.33)$$

Where μ_S is the dynamic viscosity of the continuous phase, M is the ratio between the dynamic viscosity of the dispersed phase and that of the continuous phase, and ψ is the volume fraction of the dispersed solvent phase.

The hydrophilic-lipophilic deviation (HLD) approach consists of a predictive method to establish the stability of an emulsified formulated product, with respect to both temperature and composition disturbances. It has been originally developed, in the oil and gas industry, as a system to describe and identify the range of formation of microemulsified systems. Negative HLD values suggest the formation of oil-in-water emulsions, positive values suggest the formation of water-in-oil emulsions, while HLD values in the proximity of zero indicated the formation of a three-liquid-phase system, where a stable emulsion cannot be formed.

This method consists of a correlation, based on a huge number of experimental data, considering the effect of several variables such as the presence of electrolytes, the nature of the oil as well as of the water solvent phases, the temperature, *etc.* on the emulsion type, and two different correlations have been proposed by Salager (2001), for non-ionic and ionic surfactants and they are reported in equations 3.34 and 3.35, respectively.

$$HLD = \alpha - EON + b \cdot S - K \cdot ACN - \varphi(A) + c_T \Delta T \qquad (3.34)$$

$$HLD = \sigma + \ln(S) - K \cdot ACN - f(A) + +a_T \Delta T \qquad (3.35)$$

Where α, EON and σ are model parameters that can be related, thanks to group-contribution like correlations, to the molecular structure of the surfactants, S is the salinity, ECN is the effective carbon number, and it is related to the molecular structure of the organic

solvent phase, A refers to the alcohol concentration and a_T and c_T are model parameters defining the temperature dependence.

This method does not have any actual thermodynamic but it is very useful for qualitative predictions when experimental data are not available. However, such a correlation is characterized by a high uncertainty, and it is sometimes found lacking in accuracy. Therefore, a method with thermodynamic fundaments has been proposed in this work, and details are found in §3.5.4.

Finally, also the surface tension needs dedicated models when referring to emulsified formulated products. In this context, the surface tension of the mixture is the surface tension of the continuous phase, and, in case of oil-in-water emulsions, the surface tension of the aqueous phase is usually defined as the surface tension of the aqueous surfactant solution at the critical micelle concentration.

Hence, if the organic solvent phase is the continuous phase, the model reported in §3.3.3 is used, while when the aqueous solvent phase is the continuous phase, a dedicated model is necessary. However, similarly to what mentioned in relation to the cloud point and to the critical micelle concentration, the surface tension of an aqueous surfactant solution can be modelled as a function of the only surfactant molecular structure since all other variables, that is, pressure, temperature, and composition, are fixed. Therefore, predictive methods can be applied, and the QSPR model proposed by Wang et al (2002), reported in equation 3.36, is often used.

$$\sigma = 11.98 + 0.478 \cdot nO + 0.5848 \cdot KHO - 0.0007763 \cdot E_T - 0.01053 \cdot \qquad (3.36)$$
$$\Delta H^{form} + 0.09734 \cdot D - 0.1345 \cdot nO \cdot KHO$$

Where nO, KHO, E_T and D are the model descriptors.

As for the above mentioned surfactant-related properties that can be modelled as pure component properties, group-contribution methods are considered as a possible alternative to the QSPR. However, the scarce availability of experimental data relative to the surface tension of surfactant aqueous solutions represents an obstacle to the development of such a model, for the moment.

3.5 Property Models Development

Property models of surfactants and surfactant mixtures are scarcely available in the literature, and the available models are often lacking in accuracy or non predictive in nature. Because of these reasons, an important part of this work has consisted in developing property models to be applied in the integrated methodology for the design of emulsified formulated products. Group-contribution models, based on the Marrero and Gani method have been developed in relation to the cloud point (§3.5.1) and the critical micelle concentration (§3.5.2), while a thermodynamic-based approach for the assessment of the stability of an emulsified solvent mixture has been proposed (§3.5.3).

3.5.1 Cloud Point of Non-Ionic Surfactants

An original data-set consisting of 86 nonionic surfactants have been collected from different sources (Guo et al, 2012; Ghasemi and Ahmadi, 2007; Ren et al., 2011; Rosen, 1984). The data set contains linear alkyl, branched alkyl, alkyl phenyl ethoxylates, carbohydrate-derivative ethoxylates, alkyl polyoxyethylene-polyoxypropylene copolymers and ethoxylated amides. All experimental data are measured by visual observation method in 1% aqueous surfactant solutions and they are reported in Appendix A, divided in different classes.

Before applying the Marrero and Gani CG-method to the data-set chosen for the parameter estimation step, it is necessary to analyse the matrix of group occurrences to make sure that each groups describes at least two of the surfactants presents in the data-set. A single occurrence would actually distort the performance of the model, leading to a perfect match for the compounds with those groups, providing uncertain extrapolation capabilities. Moreover, some of the experimental data are excluded from the data-set since their experimental value for the cloud point is inconsistent with other values and they are therefore identified as outliers. The outliers are identified as they are inconsistent with the assumption that the cloud point of linear alkyl ethoxylates increases with increasing length of the ethoxylated chain and with decreasing length of the carbon chain. These surfactants whose cloud point values are excluded are highlighted in grey in Appendix A.
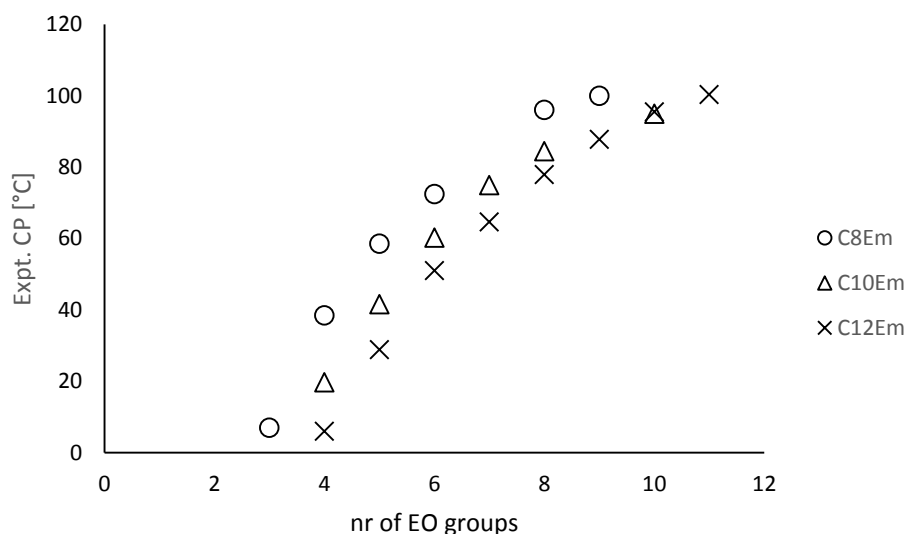
Figure 3.2 Dependence of the experimental cloud point (in °C) with the number of ethylene oxide groups in the hydrophilic head of some surfactants
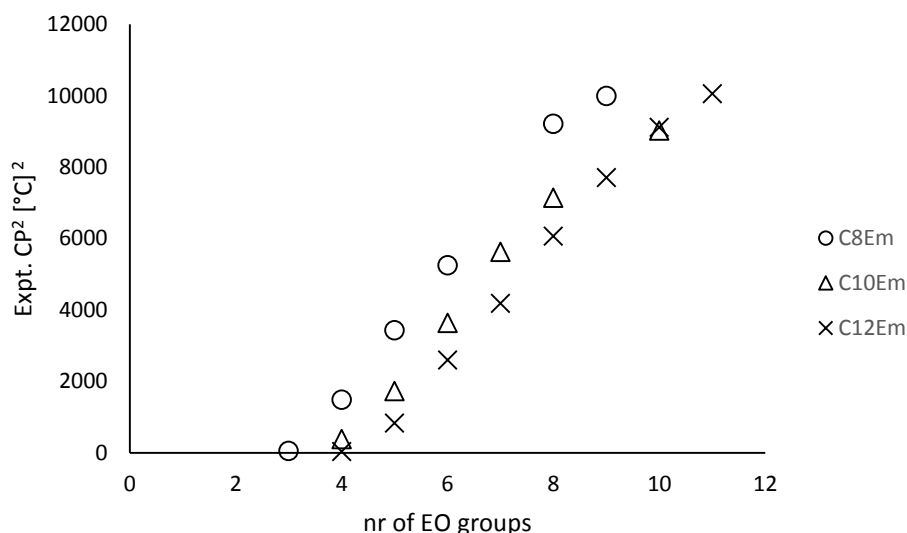
Figure 3.3 Dependence of the square of the experimental cloud point (in °C) with the number of ethylene oxide groups in the hydrophilic head of some surfactants

In order to determine the most suitable form of f(X) of the constitutive equation of the Marrero and Gani method, it is necessary to observe the trend of the experimental data of the property to be estimated as a function of the main representative groups of the chemicals under investigation. Considering the largest family of nonionic surfactants: the linear alkyl ethoxylates, the trend of the cloud point as a function of the number of ethoxylate groups (CH_2CH_2O) in the hydrophilic chain is analysed, as shown in Figures 3.2 and 3.3. As seen in Figure 3.2, the dependence of the cloud point on the number of ethoxylate groups of linear alkyl ethoxylates is not linear. On the other hand, Figure 3.3 shows that the dependence of the square of the cloud point is linear. This justifies then the choice of the form of f(X), as in the following Equation 10:

$$CP^2 = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k \qquad (3.37)$$

Where the cloud point is expressed in K. Given the equation above, in order to represent the remaining 72 compounds, 13 first order groups and 1 second order group are needed, according to the original set of parameters by Marrero and Gani. The results of the parameter estimation step performed through the step-wise regression method are illustrated in the parity plot of Figure 3.4.

The results in Figure 3.4 indicate that the accuracy of the Marrero and Gani GC-methods using only first and second order groups is not satisfactory, as quantified in Table 3.5.

Figure 3.4 Parity plot relative to the correlation of 72 data-points regarding cloud point (in °C) of non-ionic surfactants using the Marrero and Gani group-contribution method with only first and second order groups

Table 3.5 Statistical indices of performances relative to the correlation of 72 data-points regarding the cloud point of nonionic surfactants using Marrero and Gani method with only first and second order groups.

Nonionic surfactant class	SD	AAD	AAD_{max}
Linear alkyl ethoxylates	7.22	7.23	25.41
Branched alkyl ethoxylates	10.80	8.98	18.45
Phenyl alkyl ethoxylates	8.69	7.32	15.17
Alkyl polyoxyethylene-polypropylene copolymers	6.84	5.18	5.30
Carbohydrate-derivate ethoxylates	15.60	11.24	25.90
Alkyl branched ethoxylates	2.81	7.45	10.91
Overall	8.91	7.65	25.90

Where SD is the standard deviation, and AAD is the average absolute deviation.
In particular, the maximum absolute errors (column *AAD_{max}*) are too high for many categories of surfactants considered to consider the model reliable enough to be implemented in the design methodology. It is therefore considered necessary to include new third order groups in the set of parameters, in order to improve the performances of the method, as described by Hukkerikar et al. (2013), in particular relation to those compounds for which

the correlation indices were poor: branched alkyl ethoxylates and carbohydrate-derivate ethoxylates primarily. According to Marrero and Gani, in fact, the second order groups are strictly defined and one cannot arbitrarily add new second order groups as one can do with third order groups, instead.

Hukkerikar et al. (2013) recommend then to perform a step-by-step systematic data-error analysis has been performed, in order to improve the accuracy of the group-contribution method. It consists of a list of recommendations:

- o Add new data, if available, to allow a more comprehensive coverage of the molecule types;
- o Check for consistency-uncertainty of the data;
- o Identify those compounds characterized by the largest correlation errors;
- o Analyze the group descriptions of the above mentioned compounds;
- o Identify opportunities for introduction of unique third order groups for specific classes of components;
- o Regress the new group contributions.

In this work, the size of the data-set is considered satisfactory for describing the entire family of nonionic surfactants. Regarding the second point, since the source of the experimental data did not provide any information on measurement uncertainty, this information could not be included.

In this work, the following analysis has been performed:

- o Estimation of cloud point for all the compounds present in the data-set and analysis of the differences between the experimental values and the calculated ones, in order to identify compounds with the largest prediction errors;
- o Analysis of the molecular structures of those "problematic" compounds identified in the previous step and inclusion of new unique third order groups in the group-contribution model to improve the prediction performance through better correlation of data;
- o Parameter estimation after the inclusion of new third order groups in the group-contribution model to obtain the relative contributions.

The step-by-step procedure is repeated until statistical indices of performance such as the coefficient of determination (R^2), the standard deviation (SD) and the average absolute deviation (AAD) are obtained. The newly included third order groups and the relative parameters are then considered final.

The first step has been addressed by calculating the cloud point for all the compounds present in the data-set with the original set of groups of Marrero and Gani method. The statistical indices of performance are then calculated for each compound and the results, divided per surfactant class in order to identify the chemicals with the largest prediction errors, are given in Table 3.5. In general, the model needs improvements for almost all families, the largest errors being for carbohydrate-derivate ethoxylates and the lowest errors for alkyl polyoxyethylene-polypropylene copolymers. As a solution, a set of unique

and new third order groups have been added to balance these prediction errors, improving the overall performance of the model.

The definition of new third-order groups is based on the "similarity criteria" approach. It is based on the comparison of the molecular structures of the compound characterized by high prediction errors, in order to identify a set of molecules which are "similar" in nature. In order to be "similar", two or more compounds need to have one or more consecutive first order groups in common. For example, all the linear alkyl ethoxylates are "similar" since all of them have at least one CH_2 group consecutive to a CH_2O group in their structure. According to this criterion, then, "similar" compounds may be collected together and the third order group characterizing those similar molecules can be defined. On the basis of the above mentioned analysis, 5 new third order groups have been defined. Once the new set of groups has been identified, a final parameter regression is performed, where all the group contributions are estimated simultaneously. The results of this parameter regression are reported in Figure 3.5 and Table 3.6.
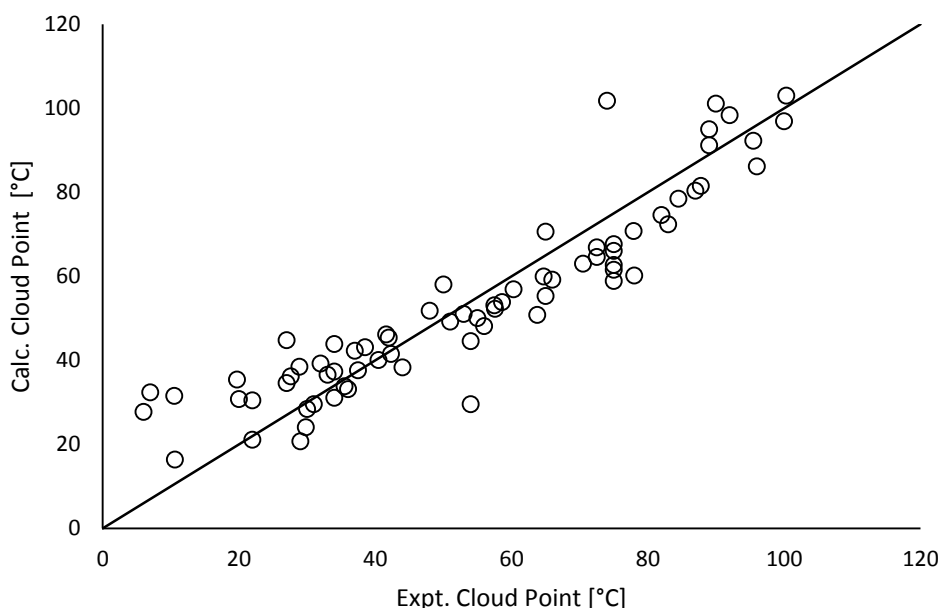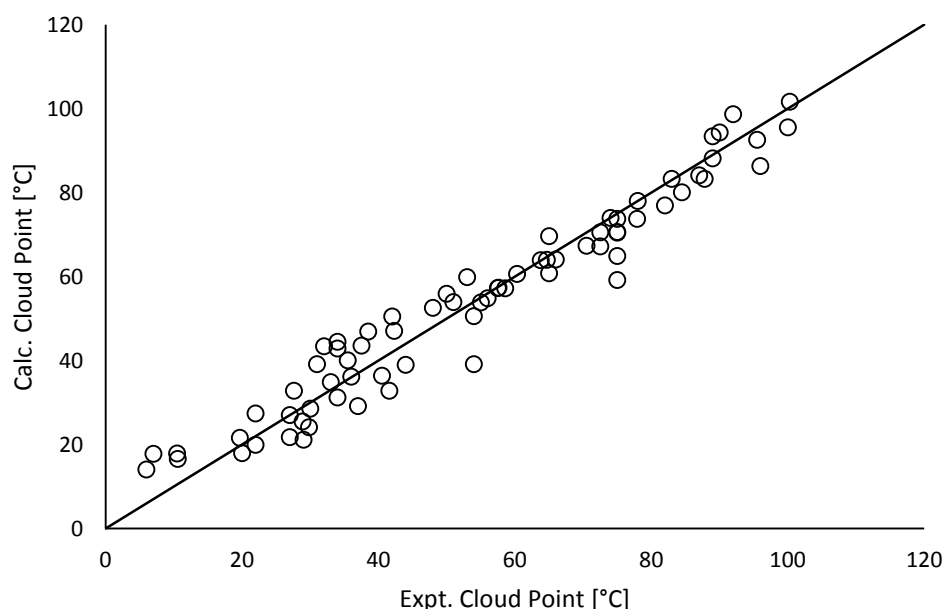

Figure 3.5 Parity plot relative to the correlation of 72 data-points regarding cloud point (in °C) of non-ionic surfactants using the Marrero and Gani group-contribution method after the addition of third order groups

Once the new set of groups has been identified, a final parameter regression is performed, where all the group contributions are estimated simultaneously. Obviously, when this approach is chosen, the absolute values of the third order contribution might be comparable or even exceed those of the first and second order groups.

Table 3.6 Statistical indices of performances relative to the correlation of 72 data-points regarding the cloud point of nonionic surfactants using Marrero and Gani method after the addition of third order groups

Nonionic surfactant class	SD	AAD	AAD$_{max}$
Linear alkyl ethoxylates	5.44	4.46	15.82
Branched alkyl ethoxylates	3.09	1.88	6.72
Phenyl alkyl ethoxylates	7.27	5.75	15.83
Alkyl polyoxyethylene-polypropylene copolymers	6.84	5.82	5.96
Carbohydrate-derivate ethoxylates	6.08	6.59	8.11
Alkyl branched ethoxylates	2.37	6.07	8.91
Overall	5.65	4.62	15.83

By comparing the results before and after the addition of the new dedicated third order groups, it can be seen that evident improvements have been achieved. In particular, after the addition of the third order groups, linear and branched alkyl ethoxylates show improved statistical indices, and in general the absolute errors have been strongly reduced. These results represent an improvement also if compared with those obtained with different QSPR methods as shown in Table 3.7.

The availability of more reliable experimental data regarding cloud points or nonionic surfactants belonging to other families (such as alkanediols, ethers, esters and fluorinated linear ethoxylates) will broaden the application range of the model. However, it can already be safely applied in the surfactant design methodology considering the limited maximum error, with basic limitation the molecular structures available.

Table 3.7 Statistical indices of performances relative to the correlation of data-points regarding the cloud point of nonionic surfactants using Marrero and Gani method before and after the addition of third order groups, compared with 3 different QSPR

Model	Data-points for the regression	SD	AAD	AAD$_{max}$
This work method without 3rd order groups	73	8.91	7.65	25.90
This work method with 3rd order groups	73	5.65	4.62	15.83
QSPR model (Guo et al., 2012)	81	9.31	7.09	50.2
QSPR model (Ghasemi and Ahmadi, 2007)	68	5.89	4.69	17.98
QSPR model (Ren et al., 2011)	78	7.46	3.13	52.8

3.5.2 Critical Micelle Concentration of Non-Ionic Surfactants

The data-set used in the present work contains 161 nonionic surfactants, covering the most common families of nonionic surfactants: linear, branched and phenyl alkyl ethoxylates; alkanediols; alkyl mono and disaccharide ethers and esters: ethoxylated alkyl amines and amides; fluorinated linear ethoxylates and amides; polyglycerol esters and carbohydrate derivate ethers, esters and thiols. In this paper, we have classified the data in terms of molecular description of each of the nonionic surfactants divided into different classes, as given in Appendix B. This classification becomes particularly important during the introduction of the unique third order groups.

When developing a QSPR model, the data-set is often divided into training and validation sets, while this is not necessary the case for group-contribution models, since the formation of a randomly selected validation set may exclude some of the model parameters, limiting the application range of the models itself. Moreover, considering as many data as possible to regress the parameters of the group-contribution model results in lower uncertainties of the estimated model parameters and consequently lower uncertainties (and better reliability) of the predicted property values (Hukkerikar et al., 2012). Therefore all the available measurements from Katritzky et al. (2008) have been considered eligible to be used in the parameter estimation step.

The definition of f(X) is specific for each property X and the selection of the most appropriate form is done by analyzing the behavior of certain class of pure compounds as their carbon number increase. For critical micelle concentration, the most suitable definition of f(X) is:

$$-log(CMC) = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k \tag{3.38}$$

In fact, by plotting different values of critical micelle concentration as a function of the number of carbon atoms in the hydrophobic chain of the surfactants, a trend as shown in Figure 3.6 is obtained.
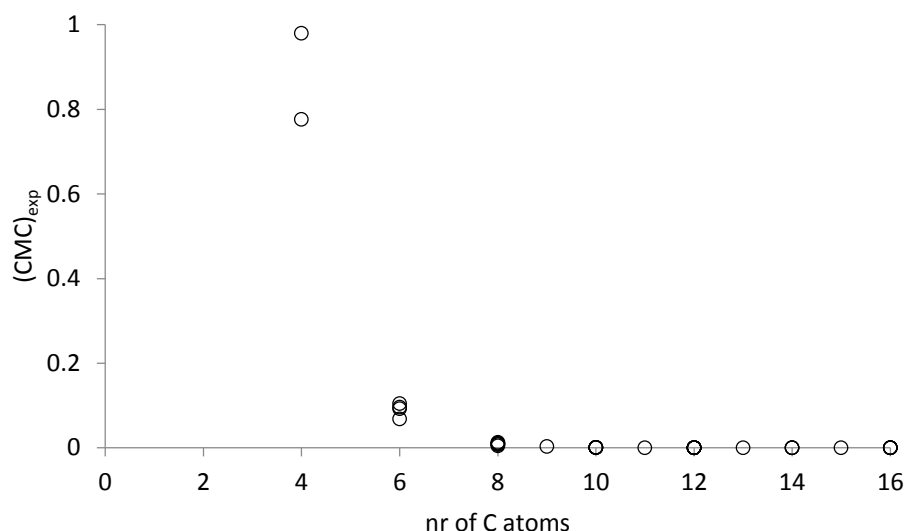
Figure 3.6 Dependence of the experimental critical micelle concentration with the number of carbon atoms in the hydrophobic tail of linear ethoxylates
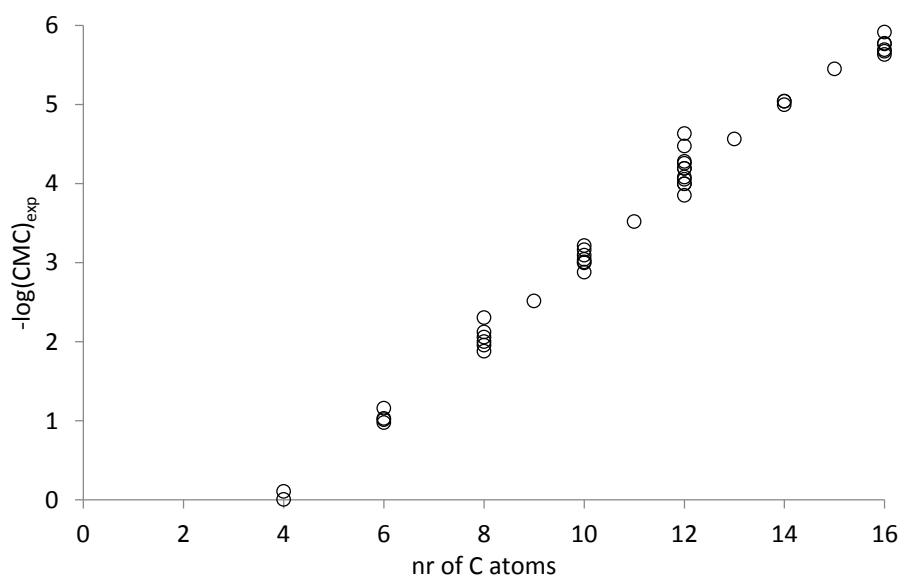
Figure 3.7 Dependence of the logarithm of the experimental critical micelle concentration with the number of carbon atoms in the hydrophobic tail of linear ethoxylates

As seen in Figure 3.6, the dependence of the critical micelle concentration of linear alkyl ethoxylates on the number of carbon atoms of the hydrophobic chain is not linear. If, on the other hand, the same analysis is carried out by plotting the logarithmic values of the critical micelle concentration (the negative sign is taken in order to obtain positive data) against the number of carbon atoms in the carbon chain of the surfactant, a trend as in Figure 3.7 is obtained. As shown in Figure 3.7, the dependence of the logarithm of the critical micelle concentration on the number of carbon atoms of the hydrophobic tail is linear. This justifies the choice of the form of f(X) as in equation (3.38).

Before the parameter estimation step, it is necessary to analyze the matrix of group oc-currences. Nine nonionic surfactants have been removed from the data-set, because of single group occurrence, and also, two surfactants have been removed from the data-set because they are isomers and their group descriptions with Marrero and Gani method are identical. Therefore, since it is not appropriate to choose one of the two values for the same group description, both data are excluded from the parameter estimation step. Therefore, the data-set originally containing 161 nonionic surfactants is reduced to 150 for the parameter regression step. To represent these compounds, 30 first order groups and 11 second order groups are needed. The results of the parameter estimation step per-formed through the step-wise regression method are illustrated in terms of a parity plot in Figure 3.8 and given in terms of statistical indices in Table 3.8.

Figure 3.8 Parity plot relative to the correlation of 150 data-points regarding critical micelle concentration of non-ionic surfactants using the Marrero and Gani group-contribution method with only first and second order groups

Table 3.8 Statistical indices of performances relative to the correlation of 150 data-points regarding the critical micelle concentration of nonionic surfactants using Marrero and Gani method with only first and second order groups

Nonionic surfactant class	SD	AAD	AAD$_{max}$
Linear alkyl ethoxylates	0.1641	0.3646	0.9106
Phenol alkyl ethoxylates	0.2096	0.2862	0.7066
Branched alkyl ethoxylates	0.0632	0.0669	0.1337
Alkanediols	0.1354	0.2946	0.5361
Alkyl mono and disaccharide ethers and esters	0.2272	0.3089	0.5382
Ethoxylated alkyl amines and amides	0.1445	0.1728	0.3496
Fluorinated linear ethoxylates and amides	0.2353	0.1787	0.4977
Polyglycerol esters	0.6783	0.5912	1.3378
Carbohydrate derivate ethers, esters and thiols	0.4704	0.3312	1.5082

The results in Figure 3.8 and Table 3.8 indicate that the accuracy of the Marrero and Gani group-contribution model using only first and second order groups is rather low compared with the performances of the same method for other properties. A step-by-step systematic data-error analysis as in §3.5.1 has been performed, and 15 new third order groups are defined, according to the similarity criteria.

Figure 3.9 Parity plot relative to the correlation of 150 data-points regarding critical micelle concentration of non-ionic surfactants using the Marrero and Gani group-contribution method after the addition of third order groups

Once the new set of groups has been identified, a final parameter regression is performed, where all the group contributions are estimated simultaneously. A relative parity plot is presented in Figure 3.9.

The performance statistics for the developed model, overall and divided per class, compared to those relative to the same model before the addition of dedicated third order groups are summarized in Table 3.9. Table 3.10 provides also a comparison of the developed GC-based method with two different QSPR-based models for the estimation of the critical micelle concentration. The relative statistical indices have been calculated based on the modeled data reported by Katritzky et al. (2008).

Table 3.9 Statistical indices of performances relative to the correlation of 150 data-points regarding the critical micelle concentration of nonionic surfactants using Marrero and Gani method after the addition of third order groups

Nonionic surfactant class	SD	AAD	AAD_{max}
Linear alkyl ethoxylates	0.1403	0.1232	0.3811
Phenol alkyl ethoxylates	0.1455	0.1638	0.3491
Branched alkyl ethoxylates	0.0727	0.0814	0.1628
Alkanediols	0.1309	0.0897	0.1676
Alkyl mono and disaccharide ethers and esters	0.0636	0.1830	0.4229
Ethoxylated alkyl amines and amides	0.1047	0.1176	0.2933
Fluorinated linear ethoxylates and amides	0.1881	0.1420	0.3569
Polyglycerol esters	0.0785	0.0781	0.1735
Carbohydrate derivate ethers, esters and thiols	0.2034	0.1250	0.6688

It can be observed that the statistical indices are much improved for each class of nonionic surfactants. In particular considerable improvements have been achieved for polyglycerol esters, compared to when only first and second order groups were used. To be fair to the authors of the QSPR-based models, we acknowledge that while ours is a strictly correlation error, theirs may or may not be correlation error. Some prediction errors which are still observed in some cases can be attributed to the insufficient availability of data for these nonionic surfactant families. When data is lacking, the "similarity" criteria cannot be applied to all those compounds with large errors, and thus unique third order groups cannot be defined. A future improvement is the inclusion in the data-set of a more compounds. That is, when more data are available, it will allow the definition of more third order groups for the remaining surfactants with large prediction errors.

Table 3.10 Statistical indices of performances relative to the correlation
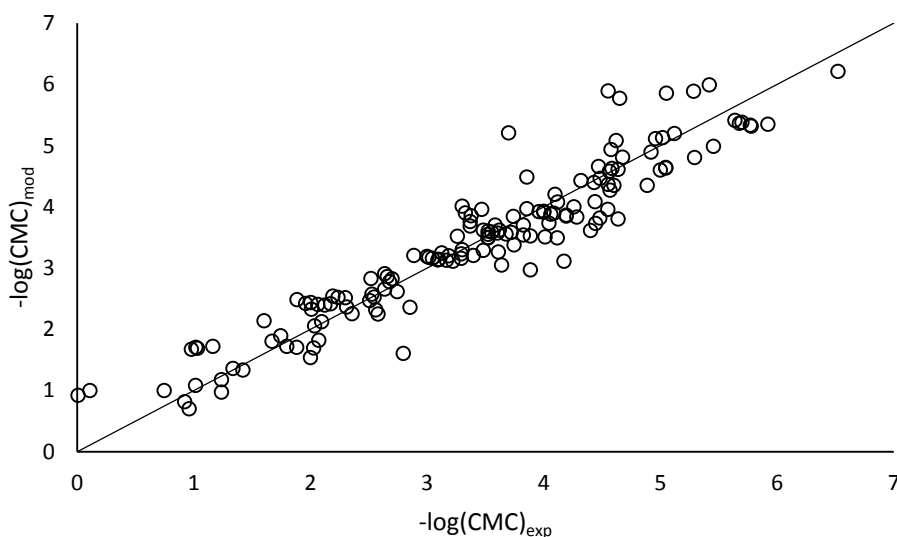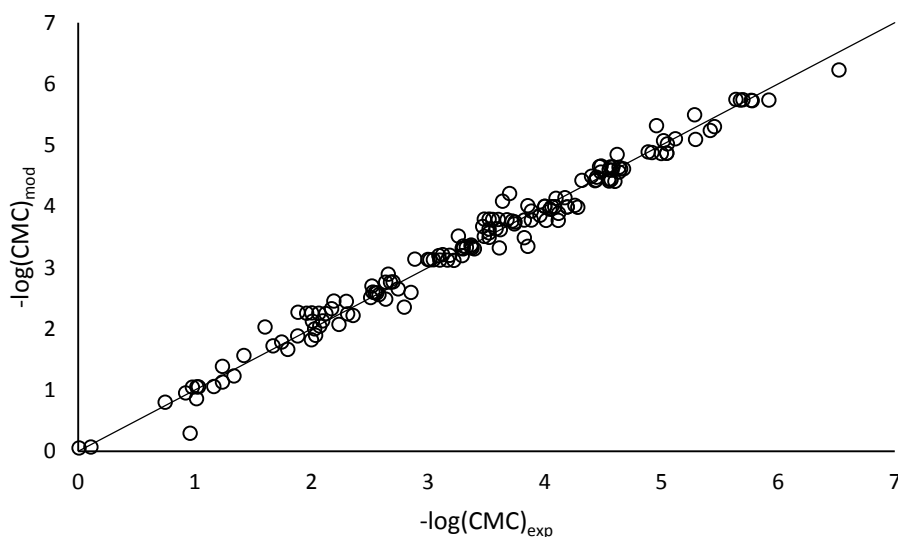of data-points regarding the critical micelle concentration of nonionic
surfactants using Marrero and Gani method before and after
the addition of third order groups, compared with 2 different QSPR

Model	Data-points for the regression	SD	AAD	AAD_{max}
This work method without 3rd order groups	150	0.41	0.31	1.51
This work method with 3rd order groups	150	0.17	0.13	0.67
QSPR model (Katritzky et al., 2007)	161	0.42	0.36	1.65
QSPR model (Katritzky et al., 2008)	161	0.29	0.25	0.89

3.5.3 Thermodynamic-base Approach for Emulsion Stability

The description of binary (water-surfactant and oil-surfactant) and ternary (water-oil-surfactant) phase behaviours with a thermodynamic model, instead of the adoption of simplified heuristic, is considered as a major progress to be implemented in the product design methodology. This way, it is possible to identify boundaries in terms of temperature and, especially, composition for a surfactant to generate a stable emulsified formulation. This paragraph will highlight some perspectives for future development of the above mentioned analysis and a few preliminary results; to this point, however, the application of surfactant-related phase behaviours is limited to the availability of experimental data relative to the systems of interest. Therefore this approach can be applied for product analysis as well as for verification of the design obtained from the methodology.

Binary Systems: Water-Surfactant and Oil-Surfactant

Water-surfactant phase diagrams are fundamental when the surfactant is expected to be mainly dissolved in the water-phase (high values of the hydrophilic-lipophilic balance), which leads to the formulation of oil-in-water emulsions. On the other hand, the phase diagrams between oil and surfactants are to be considered primarily when a water-in-oil emulsion is desired.

As an example of water-surfactant phase behaviours, Figure 3.1 presented the phase diagram of the system between water and dodecyl-esaethylene oxide. For emulsion-based chemical product design, therefore, it is relevant to define the boundaries in terms of temperature and concentration, so that the designed formulation lies in the area defined as L_1, where the models presented in this chapter can be safely applied. Mitchell et al. (1983) and Sjöblom et al. (1987) provide a satisfactory amount of experimental data as well as theoretical explanations for the formation of the different meso-phases, relatively to aqueous surfactant solutions. The methods proposed, however, cannot be used for the prediction of these phase boundaries and therefore they cannot be applied in the framework for chemical product design. Approximate predictions of these phase behaviours, based on

the analysis of several phase diagrams in parallel with the molecular structure of the species involved, are however expected to be possible by the authors and they are considered to be a potential important development on the way for a fully model-based methodology for emulsion-based chemical product design. The adoption of such correlations can lead to the calculation of simplified water-surfactant phase behaviours as described in Figure 3.10, in relation to the system reported in Figure 3.1.
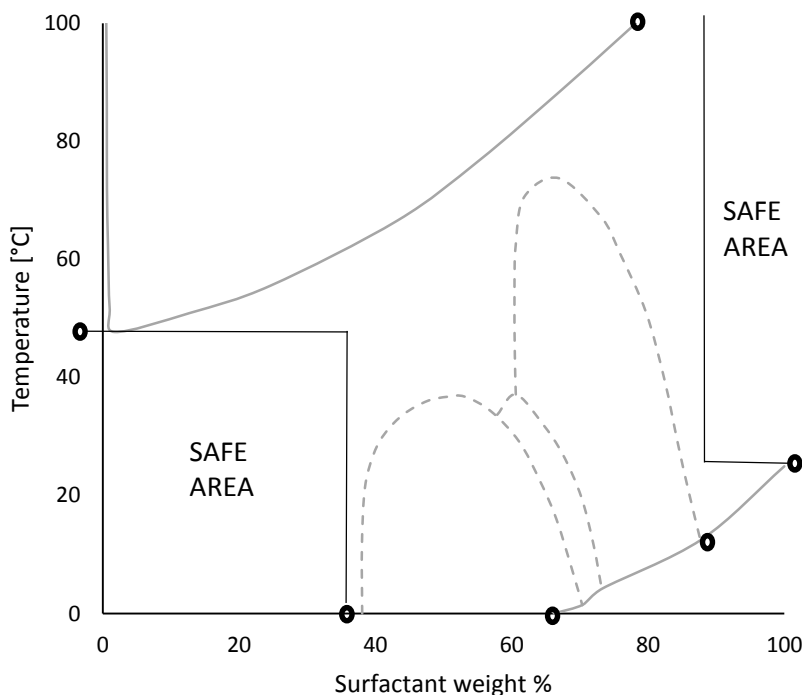
Figure 3.10 Simplified phase envelope of an aqueous mixture of a polyoxyethylene-based nonionic surfactants. Data are taken from Mitchell et al. (1983).

When a water-in-oil emulsion is wanted, on the other hand, the phase behaviour of surfactant mixtures with oil is relevant. Figure 3.11 shows examples of the phase behaviours between oils and surfactants, in terms of liquid-liquid miscibility boundaries. The reference surfactant is hexyl-pentaethylene oxide, while four different alkanes are considered as the oil-phase.

In Figure 3.11, only the miscibility curve as function of surfactant concentration and temperature is reported, while the formation of micellar solutions of standard appearance or of viscous meso-phases with different self-assemblies of the surfactant in absence of water is debatable (Ray, 1971). For use in emulsion-based chemical product design, it is necessary that the designed formulation lies above the line of the miscibility gap. It is easy to determine in Figure 3.11 a trend of the miscibility curves as a function of the number of carbon atoms of the *n*-alkane considered. This leads to the consideration that a correlation based on the molecular structure of the chemicals involved in the phase equilibrium can approximately describe these curves and therefore a correlative model

may be developed and applied for use in emulsion-based chemical product design.
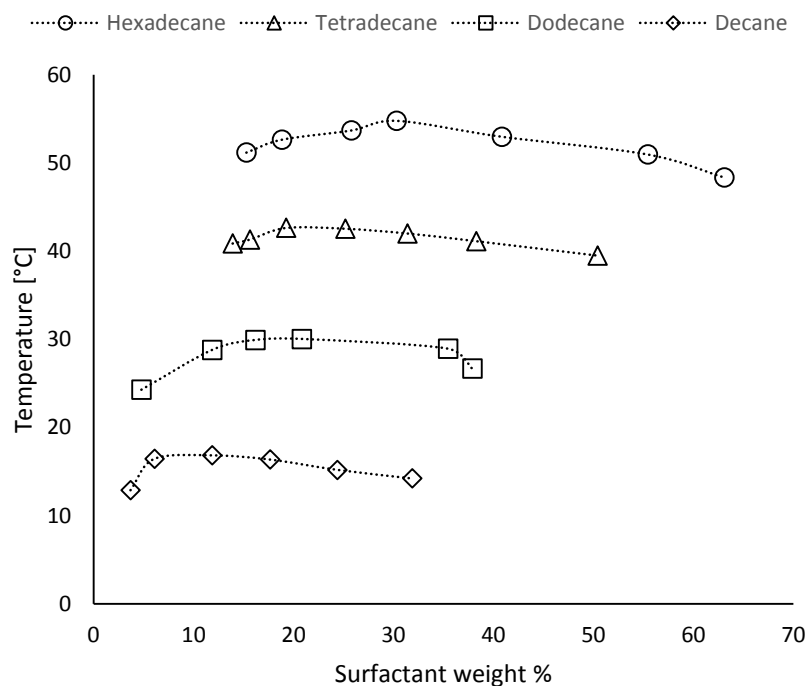
Figure 3.11 Phase envelope of the mixtures between hexyl-pentaethylene oxide and four different alkanes. Data are taken from Kahlweit and Strey (1985).

In relation to both water-surfactant and oil-surfactant phase behaviours, however, experimental measurements are needed in order to define numerical boundaries on the composition of the desired formulation, since model-based generation of data is not considered yet reliable.

Ternary Systems: Water-Oil-Surfactant

The understanding of the behaviour of ternary water-oil-surfactant systems is also considered to be crucial, in order to determine temperature and composition boundaries for a stable emulsion. These type of phase envelops can be represented in several ways, since many variables are involved; in relation to emulsion-based chemical product design, the most useful alternative is represented by the use of the so-called Kahlweit's fish phase diagram (Kahlweit and Strey, 1985). Here, ternary water-oil-surfactant data are drawn in an X-Y diagram, where the surfactant content (usually in weight percentage) is in the X-axis, while the temperature is on the Y-axis. These diagrams represent a valid tool for emulsion-based product design since different types of products can be recognized and the possibilities for the formation of each of them are easily identified, given the temperature and the composition of the formulation. An example of such as "fish-type phase diagram" for the system water-tetradecane-2-butoxyethanol is shown in Figure 3.12.
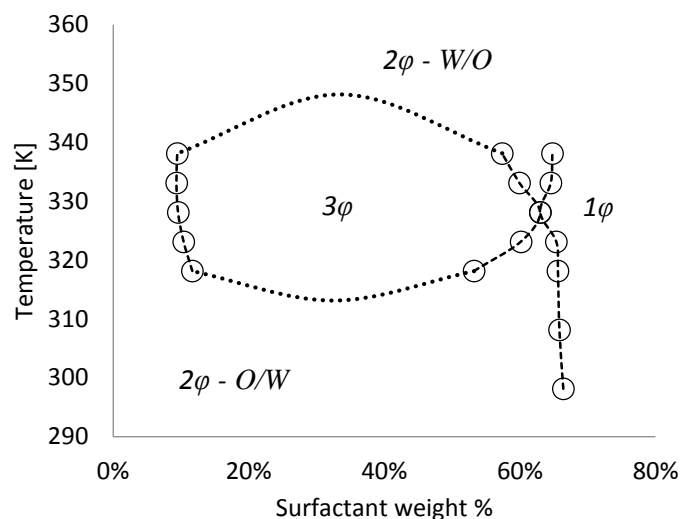
Figure 3.12 Calculated fish-diagram of the system water-tetradecane-2-butoxyethanol; water-oil ratio: 7.03. Data are taken from Lin and Chen (2004).

In Figure 3.12 different areas are can be identified: the region defined with the symbol 1φ represents the area where a micro-emulsion can be formed, the region defined with the symbol 2φ (both 2φ – W/O and 2φ – O/W) is the area where an emulsified product can be formulated, while the region identified by the symbol 3φ is a hybrid domain, where an emulsion and a microemulsion may coexist. The emulsion domain (2φ) consists of two areas: one above the hybrid domain, described by the symbol 2φ – W/O, where a water-in-oil emulsion can be formed and another below the hybrid domain, described by the symbol 2φ – O/W, where an oil-in-water emulsion is favoured instead. Consequently, the region of the emulsion domain located at intermediate temperatures between the two above mentioned areas represents an unstable region where it is not recommended to design an emulsified product, since its life time is expected to be limited. As a part of an emulsion-based chemical product design procedure, then, it is necessary to make sure that the designed formulation lies in the 2φ – W/O area when a water-in-oil emulsion is desired and in the 2φ – O/W if an oil-in-water emulsion is wanted. This type of diagrams can be used both during the design of the surfactant and during the verification of the designed product. The authors are not aware of any reliable model for the prediction of such phase equilibria, and therefore this analysis can be performed up to now only when experimental data are available.

When predictions are necessary, the hydrophilic-lipophilic deviation approach is applied instead. With this method, if the calculated HLD-value is zero, then the formulation is located in the 3φ domain of Figure 3.12 and therefore an unstable system is expected. On the other hand, if a positive value is obtained, then a water-in-oil emulsion is favoured, while if a surfactant is characterized by a negative value of its HLD, then an oil-in-water emulsion may be formed. The higher the absolute value of the HLD of the surfactant is, the more stable the emulsion formed is expected to be, since it is located further away from the unstable region identified by the hybrid domain. This method does not have the

thermodynamic basis of the representation of the ternary phase diagram, but it can be used as a qualitative predictive model when the needed experimental data are not available.

4

METHODS
AND TOOLS

This chapter focuses on the methods and tools which have been developed in this work, for a reliable solution of emulsified product design problems. The methods consist of the algorithm for decision-making regarding the solvent mixture design, fundamental for defining the overall formulation composition, while the tools here are the structured databases, the systematic knowledge-base, and the computer program. As it has already been highlighted in Chapter 3, the property models, which are to be considered tools themselves, have been dedicated a whole chapter for their importance in this work.

In §4.1, the knowledge-base and its structure is highlighted, while the structured databases are presented, with links to the property models that have been applied for their development §4.2. In §4.3, the algorithm for the mixture design in the emulsified form is given, and finally, in §4.4, the software tools employed in this work are presented. First some of the software used for miscellaneous calculations are reported, and then the Virtual Process-Product Design Laboratory is introduced. This represents the extension of a previously developed software for the computer-aided solution of problems relative to chemical product design (Conte, Gani and Malik, 2010), so that emulsified formulations can also be designed, according to the methodology presented in this work.

4.1 Knowledge-base

The formulation and the solution of a chemical product design problem requires the collection and further availability of a large number of data and information, coming from different sources and specific to different disciplines: expertise in chemistry, engineering, marketing, manufacturing and economics are in fact necessary (Cheng et al., 2009).

In particular, the input information required by the design methodology (indifferently from its nature: model-based, experiment-based or integrated) such as the customer needs, the product quality factors, the product technical specifications, *etc.* are obtained by means of different sources (market and customer surveys, patent, literature, *etc.*) and its collection and provision has been rarely systematized.

In this work, a systematic knowledge-base has been developed, so that the necessary information and data obtained for the development of the case studies presented in Chapter 6 is stored and provided, when needed, in an efficient way. Moreover, the structure allows the integration of the above mentioned knowledge-base with the information relative to other case studies previously developed, as for example by Conte et al. (2013) and by Yunus et al. (2014), and with any future work in the field of chemical product design.

Figure 4.1 illustrates the multi-layer structure of the knowledge-base, developed in this work.
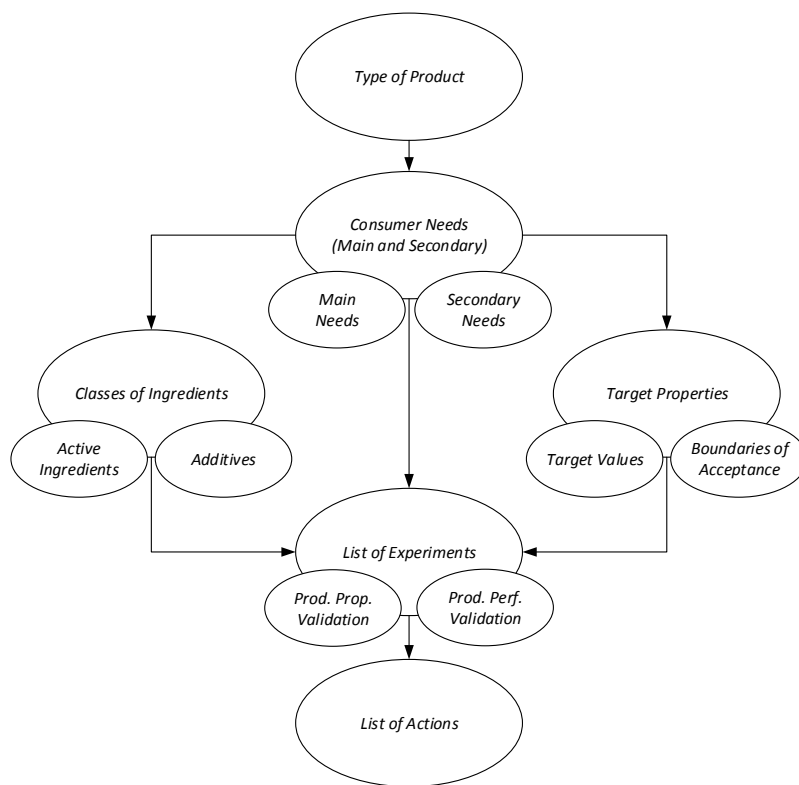
Figure 4.1 Structure of the knowledge-base
for chemical product design

The structure has been developed in order to reflect the systematic structure of the framework: first the problem definition stage, then the model-based stage for synthesis and design, and finally the experiment-based stage for validation and refinement.

The knowledge-base plays a fundamental role in particular at the level of the problem definition, by providing information about the necessary classes of ingredients and the target properties, that is, the thermo-physical properties that the formulation need to sat-

isfy. Moreover, it assists in the experiment-based stage in focusing the experimental efforts and in translating the measurements in further refinements for the candidate formulated product. The knowledge-base developed in this work is further described one layer by one, using the conceptual case-study of a mosquito repellent lotion in the emulsified form, to highlight it.

Type of Product

The first layer of the knowledge-base consists of a short list of pre-defined chemical products, all belonging to the category of the emulsified formulated products. These coincide with the case studies that have been developed in this work and that are collected in Chapter 6: a UV sunscreen, an industrial detergent, an ink-jet ink, and a hand-wash, plus the conceptual case-studies of a mosquito repellent, presented here only for explanatory reasons.

Such information represents the input of the overall integrated methodology for the design of emulsified formulated products. All the necessary information relative to these products are stored in the knowledge-base, so that as soon as one of these products is considered, these are provided in an efficient and reliable way.

Consumer Needs

The second layer of this knowledge-base consists of the consumer needs. It contains the product attributes (ψ_i) that are usually required by the specific type of product defined in the previous layer of the knowledge-base. These attributes range from sensorial factors (odor, color, *etc.*), to safety related properties (explosive limits, flash point, *etc.*), to product performances (ability to be sprayed, or spread, *etc.*).

Among the consumer needs, moreover, it is necessary to distinguish between main consumer needs (ψ_m) and secondary consumer needs (ψ_s). The main consumer needs are the main functions of the products: the main reasons for which consumers buy the product. The secondary needs, instead, are related to other performance criteria such as the form of the product (liquid, solid, *etc.*), cosmetic properties, and so on. Such a discrimination is fundamental since the definition of active ingredients and additives originates from this decision.

These information are necessary in the problem definition stage of the integrated methodology, and they have been retrieved from patents, real products, industrial experience and common sense. Table 4.1 gives an insight of the main and secondary consumer needs related to a mosquito repellent in the emulsified form. It has to be noticed, however, that in this layer it is not straightforward to identify the desired physical form of the product, as the consumer needs are the same, for example, if the product has to be designed as a homogeneous or as an emulsified formulated product (Conte et al, 2009).

Table 4.1 Knowledge base relative to the main and secondary consumer needs of a mosquito repellent in the emulsified form

Main Consumer Needs (ψ_M)	Secondary Consumer Needs (ψ_S)
Protection from mosquitos	Material compatibility
	Nice odor
	Good skin feeling
	Long durability
	Low toxicity
	High stability
	Spray-ability
	Long shelf life

Classes of Ingredients

In the third layer of this knowledge-base consists, the concept of the necessary classes of ingredients is introduced. The necessary classes of ingredients (ξ_i) are defined as those categories of chemicals that need to be added to the formulated products in order to satisfy some of the consumer needs. Some of the product attributes, in fact, can be quantified in terms of thermo-physical properties, while for others such a translation is not possible. A typical example considers cosmetic properties such as color and odor; these product attributes cannot be quantified by means of properties and/or property models, but through the addition of certain classes of chemicals it is possible to guarantee the satisfaction of the consumer needs.

Among these ingredients, moreover, it is fundamental to distinguish the active ingredients (ξ_{AI}) from the additives (ξ_{ADD}). The first class of chemicals is represented by those ingredients that satisfy the main needs (ψ_m), while the second is represented by those classes of chemicals relative to the secondary needs (ψ_s).

It is important to stress the fact that not all the consumer needs have one or more respective necessary classes of ingredients, as for some of the consumer needs, a translation into thermo-physical properties is possible.

In relation to the conceptual case study of an emulsified mosquito repellent, Table 4.2 provides the knowledge-base relative to the consumer needs.

Table 4.2 Knowledge base relative to the necessary classes of ingredients
of a mosquito repellent in the emulsified form

Consumer Need	Main or Secondary	Class of Ingredient	Active Ingredient or Additive
Protection from mosquitos	Main	Mosquito repellent	Active Ingredient
Nice odor	Secondary	Aroma	Additive

Target Properties

The fourth layer of this knowledge-base introduces the concept of the target properties is introduced. The target properties (ζ_i) consist of the direct translation of the consumer needs into thermo-physical properties, together with relative target values and boundaries of acceptance. Typical examples of such a translation are relative to the physical form of the desired product: if a liquid product is desired, in fact, the solvent mixture needs to be characterized by a melting temperature and a boiling temperature lower and higher than the room temperature, respectively.

Similarly to what mentioned in relation to the classes of ingredients, also the target properties are classified depending on the classes of ingredients the relate to. Accordingly, four categories of target properties can be identified: the target properties relative to the active ingredients (ζ_1), those relative to the additives (ζ_2), those relative to all the ingredients (ζ_3), and those relative to the overall product (ζ_4).

Table 4.3 Knowledge base relative to the target properties
of a mosquito repellent in the emulsified form.
(T_{90}: evaporation time [s], LC_{50}: lethal concentration [mol/l],
HLB: Hydrophilic-lipophilic balance, HLD: Hydrophilic-lipophilic deviation,
V_m: molar volume [l/kmol], v: kinematic viscosity [cS])

Consumer Need	Main or Secondary	Target Properties	Boundaries of Acceptance	Ingredient of Interest
Long durability	Secondary	Evaporation time	$500 < T_{90} < 1500$	Overall product
Low toxicity	Secondary	Toxicity parameter	$LC_{50} > 0.39$	All ingredients
High stability	Secondary	Solubility parameter HLB HLD	$\delta_{T,Al} - 3 < \delta_T < \delta_{T,Al} + 3$ HLB > 10 HLD ≠ 0	Overall product Surfactant Overall product
Spray-ability	Secondary	Molar volume Viscosity	$20 < V_m < 50$ v < 75	Overall product

Similarly to the classes of ingredients, the target properties usually do not relate to all the consumer needs, as sometimes a direct translation into thermo-physical properties is not feasible.

Table 4.3 illustrates the knowledge-base relative to the target properties of a mosquito repellent in the emulsified form.

List of Experiments

The above mentioned first four layers represent the part of knowledge-base assisting the problem definition stage of the integrated methodology for the design of emulsified formulated products. The fifth layer, instead, inaugurates the section of the knowledge-base dedicated to the experiment-based stage of the methodology. As illustrated in Figure 4.1, the knowledge-based relative to the list of experiments is connected with three of the previous layers: the consumer needs, the classes of ingredients and the target properties. This layer recommends a list of experiments to be performed in order to validate the product that has been designed in the model-based stage on the basis of the necessary categories of ingredients and properties generated in the problem definition stage.

Table 4.4 Knowledge base relative to the list of experiments
for a mosquito repellent in the emulsified form.
(PIT: Phase inversion temperature)

Consumer Need	Product Property Validation Experiment	Product Performance Validation Experiment
Protection from mosquitos	-	Panel test of product activity
Material compatibility	-	Panel test of fabric compatibility
Nice odor	-	Panel test for the odor
Good skin feeling	-	Measurement of pH Panel test for the skin feeling
Long durability	Measurement of T_{90}	-
Low toxicity	Measurement of the PIT	-
High stability	-	Solubility test Solvent mixture stability test Product stability test
Spray-ability	Measurement of V_m and v	Panel test of spray-ability
Long shelf life	-	Shelf life test

The recommended experiments are divided into two categories:
- The experiments validating the product properties; these experiments have the objective of measuring the target thermo-physical properties of the product, and comparing them with the predictions of the model-based stage;

- The experiments validating the product performances; these experiments have the objective of assessing the performances of the product, and comparing them with the expected performances.

Table 4.4 gives the list of experiments, together with the relative consumer needs, that are stored in the knowledge-base for to a mosquito repellent in the emulsified form.

List of Actions

The sixth layer of the knowledge-base developed in this work is strictly connected with the previous one. In fact, it is necessary to provide a list of action to be performed in case any of the experiments to be performed fail to validate either a product property, either a product performance.

Each of the measurements and tests that are in the list of experiments is accompanied with one or more actions to be taken, in case the experimental results are not satisfactory. This is highlighted in Table 4.5, in relation to an emulsified mosquito repellent.

Table 4.5 Knowledge base relative to the list of actions
for a mosquito repellent in the emulsified form.
(DSD: Droplet size distribution)

Validation Experiment	Action 1	Action 2
Measurement of T_{90}	Change the solvent mixture	-
Measurement of PIT	Change the solvent mixture	-
Measurement of V_m	Add an adequate additive	Change the solvent mixture
Measurement of v	Change the DSD	Add an adequate additive
Panel test of product activity	Change the solvent mixture	Change the active ingredient
Panel test of fabric compatibility	Change the solvent mixture	-
Panel test for the odor	Add an adequate additive	Change the additive
Measurement of pH	Add an adequate additive	Change the additive
Panel test for the skin feeling	Add an adequate additive	Change the solvent mixture
Solubility test	Change the solvent mixture	-
Solvent mixture stability test	Change the solvent mixture	-
Product stability test	Change the solvent mixture	-
Panel test of spray-ability	Change the solvent mixture	-
Shelf life test	Change the solvent mixture	-

The actions reported in Table 4.5 underline the fact that changing the solvent mixture can solve most of the problems, if some experimental results are not satisfactory. However, as changing the solvent mixture can influence the choice of the additives, it is recommended to add and/or change the additives, if possible. Moreover, it needs to be stressed that it is not always straightforward to define how to change the solvent mixture (such as, changing the water-oil ratio, changing the surfactant system, changing one of the two solvent phases, *etc.*), and sometimes different recommended actions might disagree each other. Therefore, the final choice of the refinements to be done over the candidate formulated product relies on the experience and on the common sense of the formulation engineers and chemists involved with the design problem.

4.2 Structured Databases

Chemical databases are a fundamental tool for the solution of chemical product design problems. In fact, most often, the generation of new molecules by means of computer-aided molecular design techniques is not necessary when designing a chemical formulation, both because of the complexity of the problem, and for the availability of the ingredients for experiments for validation and/or refinement. It is common practice, therefore, that the different ingredients of a formulated chemical product are selected from dedicated databases, by means of model-based techniques.

For this reason, several different databases have been created, as part of this work, collecting the categories of ingredients necessary for the solution of the case studies given in Chapter 6. The above mentioned databases, are structured on the basis of different criteria, such as, the type of ingredients (active ingredients, solvents, and additives), the activity of the ingredients (aromas, colorants, UV filters, *etc.*) and the qualitative solubility/miscibility of the ingredients.

In §4.2.1, the databases relative to the active ingredients and the additives are listed and their specific needs are described, while §4.2.2 focuses on the databases of solvents and, finally, the surfactants databases are presented in §4.2.3.

The distinction above is due to the fact that different databases have different characteristics and requirements, as these are defined in agreement with the overall methodology for the design of emulsified formulated products highlighted in Chapter 5.

4.2.1 Active Ingredients and Additives Databases

In relation to the case studies presented in Chapter 6, the following databases have been developed:

- UV-A absorbers database (Ansmann et al., 2001, Regulation (EC) No. 1223-2009 of the European Parliament and of the Council of 30th November 2009 on Cosmetic Products), contains a list of chemicals that can block the UV-A radiations by absorbing them; they are used in sunscreen products;

- UV-B absorbers database (Ansmann et al., 2001, Regulation (EC) No. 1223-2009 of the European Parliament and of the Council of 30th November 2009 on Cosmetic Products), contains a list of chemicals that can block the UV-B radiations by absorbing them; they are used in sunscreen products;
- UV filters database (Ansmann et al., 2001, Regulation (EC) No. 1223-2009 of the European Parliament and of the Council of 30th November 2009 on Cosmetic Products), contains a list of chemicals that can block the UV radiations by acting as physical screens; they are used in sunscreen products;
- Antioxidant database (Ansmann et al., 2001), contains a list of chemical that can prevent the formation of free radicals, thus protecting from skin-ageing; they are used in sunscreen and skin-care products;
- Preservatives database (Ansmann et al., 2001), contains a list of chemicals that can prevent undesirable product changes; they are used in skin-care and body-care products;
- Aromas database (Arctander, 1969), contains a list of chemicals that can mask unpleasant scents, thus enhancing the end-use product properties; they are used in the majority of the consumer products;
- Co-surfactants database (Surabhi et al., 2010), contains a list of chemicals that can improve the elasticity of the oil-water interfacial film, thus enhancing the stability of emulsions and micro-emulsions; they are used in emulsified formulated products;
- Builders database, contains a list of chemicals that can remove undesired ions by complexation or precipitation; they are used in detergency products;
- Buffering agents database, contains a list of chemicals that can prevent undesired changes of the pH of a solution; they are used in emulsified formulated products;
- Bleaching agents database, contains a list of chemicals that remove undesired chemicals by oxidation; they are used in detergency products;
- Colorants database, contains a list of chemicals that can modify the color of a solution; they are used in a wide range consumer products;
- Anti-microbial agents database, contains a list of chemicals that can kill microorganisms, or inhibit their growth; they are used in detergency products;
- Emollients database, contains a list of chemicals that can improve the moisture of the skin; they are usually employed in skin-care and body-care products.

All the above mentioned databases, however, do not consists of a mere list of chemicals, as a minimum set of information and thermo-physical properties are also necessary. These information have been retrieved from the literature or, if a dedicated pure component property model was available, predicted. Table 4.6 illustrates the minimum set of information required for the databases of active ingredients and additives.

Table 4.6 Example of list of chemicals with necessary information and
thermo-physical properties for a database of active ingredients and additives.
(MW: molar weight)

Commercial Name	Qualitative Solubility	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [˚C]	Toxicity Parameter [mol/m³]	Cost [$/kg]
Ingredient 1	Aqueous	224	1.07	550	>100	5.12	34
Ingredient 2	Organic	288	0.91	475	>100	4.24	27
…	…	…	…	…	…	…	…

It has to be underlined that in case no information about the qualitative solubility is available, a reliable alternative consists of the calculation of the Hansen (recommended) or the Hildebrand solubility parameters.

Among the active ingredients defined in the case studies of Chapter 6, also anionic as well as non-ionic surfactants can be found. However, because of the peculiar properties characterizing these categories of chemicals, their databases are presented in a dedicated paragraph.

4.2.2 Solvents Databases

In relation to the case studies presented in Chapter 6, the following databases have been developed:

- Water database;
- Water insoluble alcohols, contains a list of alcohols that are well known to be water insoluble; they are usually employed as the organic solvent phase of homogeneous and emulsified formulated products;
- Water soluble alcohols, contains a list of alcohols that are well known to be water soluble; they are usually employed as the aqueous solvent phase of homogeneous and emulsified formulated products;
- Esters database, contains a list of esters, chemicals that are well known to be water insoluble; they are usually employed as organic solvents in skin-care and body-care products;
- Vegetable oils database, contains a list of commercial vegetable oils, well known to be water insoluble; they are usually employed in several industries, as they act as the organic solvent phase of emulsified formulated products.

All these solvents above mentioned databases also require a minimum set of information, in terms of thermo-physical properties, to be of any use in this integrated methodology. Given the pure component property models listed in Chapter 3, in case any of these necessary properties cannot be retrieved from the literature, it can be predicted.

Table 4.7 highlights the minimum set of information required for the solvents databases.

Table 4.7 Example of list of chemicals with necessary
thermo-physical properties for a database of solvents.

Commercial Name	Solubility Parameter [MPa$^{1/2}$]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [°C]	Toxicity Parameter [mol/m^3]	Cost [$/kg]
Ingredient 1	4.1	224	0.94	31	>100	3.64	75
Ingredient 2	5.9	288	0.88	27	>100	4.24	82
…	…	…	…	…	…	…	…

It has to be underlined, however, that for design purposes the availability of the Hansen solubility parameters, instead of the Hildebrand solubility parameter (as in Table 4.7) is highly recommended.

4.2.3 Surfactants Databases

In relation to the case studies presented in Chapter 6, the following databases have been developed:

- Anionic surfactants database, contains a list of pure anionic surfactants; they are usually employed as active ingredients in commercial detergents and as emulsifiers in emulsified formulated products;
- Non-ionic surfactants database, contains a list of pure non-ionic surfactants; they are usually employed as active ingredients in detergents and as emulsifiers in emulsified formulated products;
- Commercial surfactants database, contains a list of commercial surfactants, often consisting of surfactant mixtures; they are usually employed as active ingredients in detergents and as emulsifiers in emulsified formulated products.

As it has already been mentioned, the surfactants are used as active ingredients for some categories of formulated products (detergents), while they are used as emulsifier almost for the totality of the emulsified formulated products in the market. Because of their peculiar chemical structure and properties, however, a distinct paragraph has been dedicated to the relative databases.

In fact, the minimum set of thermo-physical properties required by the methodology, includes some peculiar properties of surfactants such as the critical micelle concentration, the cloud point (for non-ionic surfactants), the Krafft temperature (for anionic surfactants), the surface tension reduction, and the hydrophilic-lipophilic balance. As reported in Chapter 3, all these properties can be predicted, if the experimental values are not available, with dedicated pure component property models. Table 4.8 illustrates the information stored in the surfactants databases.

Table 4.8 Example of list of chemicals with necessary
thermo-physical properties for a database of surfactants.
*(Comm. Name: commercial name, CMC: critical micelle concentration, CP: cloud
point, σ: surface tension, HLB: hydrophilic-lipophilic balance,
ρ: density, μ: liquid viscosity, T_f: flash point, LC_{50}: toxicity parameter)*

Comm. Name	CMC [mol/L]	CP [°C]	σ [mN/m]	HLB [-]	MW [g/mol]	ρ [g/ml]	μ [cP]	T_f [°C]	LC_{50} [mol/m³]	Cost [$/kg]
Surf. 1	0.014	72	29	17.2	645	1.07	370	75	3.58	175
Surf. 2	0.005	52	37	12.4	821	1.11	405	67	4.58	231
…	…	…	…	…	…	…	…	…	…	…

4.3 Emulsified Mixture Design

The mixture design problem addressed by this algorithm is a typical reverse design problem where, given a set of constraints on a set of target properties, the solvent mixtures satisfying the constraints are to be identified, and further ranked according to a specific objective function, or performance index.

In general terms, this problem is formulated as a Mixed Integer Non-Linear Programming (MINLP) problem, as given by equations (4.1) to (4.4).

$$max/min \, F_{obj}(X, Y, C, E, S, P) \tag{4.1}$$

$$f_1(X, Y) > 0 \tag{4.2}$$

$$\zeta_{i,L} < f_2(X, Y, \zeta_i) < \zeta_{i,U} \tag{4.3}$$

$$f_3(X, Y) = 0 \tag{4.4}$$

In equation (4.1), F_{obj} is the objective function to be maximized/minimized, and it consists of one or a combination of the following performance indices: the composition (X), the type of mixture (Y), the cost (C), the environmental impact (E), a safety factor (S), and the product performance (P). Equation (4.2) represents the mixture stability equations, where the mathematical expression of f_1 is given with respect to the miscibility and/or solubility condition that must be satisfied. Equation (4.3), instead, represents the mixture property equations, where $\zeta_{i,L}$ and $\zeta_{i,U}$ are the lower and the upper property boundaries of acceptance, respectively, and f_2 is the mathematical representation of the property model. Finally, equation (4.4) represents the product composition equations, where limitations in the composition of the formulated product are given; the mathematical form of f_2, then, depends on the specific limitation to be satisfied.

According to Yunus et al. (2013), the solution of the above mentioned problem, when large chemical databases are involved, often creates a combinatorial explosion. Therefore, a systematic decomposition based solution approach (Karunanithi et al., 2005) is preferred, in order to manage the complexity of the mixture design efficiently, and to further reduce the search space, by decreasing the number of feasible mixtures in a hierarchy of calculations of increasing complexity.

As an algorithm for the design of emulsified solvent mixtures, an aqueous solvent phase, an organic solvent phase, and a surfactant system are required. They are designed simultaneously in a multi-level structure, where the input information are the relative databases (with all the necessary pure component properties, as highlighted in §4.2), the constraints on the mixture target properties and the relative property models with all the necessary parameters, if needed, and the design temperature. As output, the algorithm returns the candidate mixtures satisfying the given constraints, with their composition, their mixture properties and their cost.

Figure 4.2 illustrates the multi-level structure of the proposed algorithm.
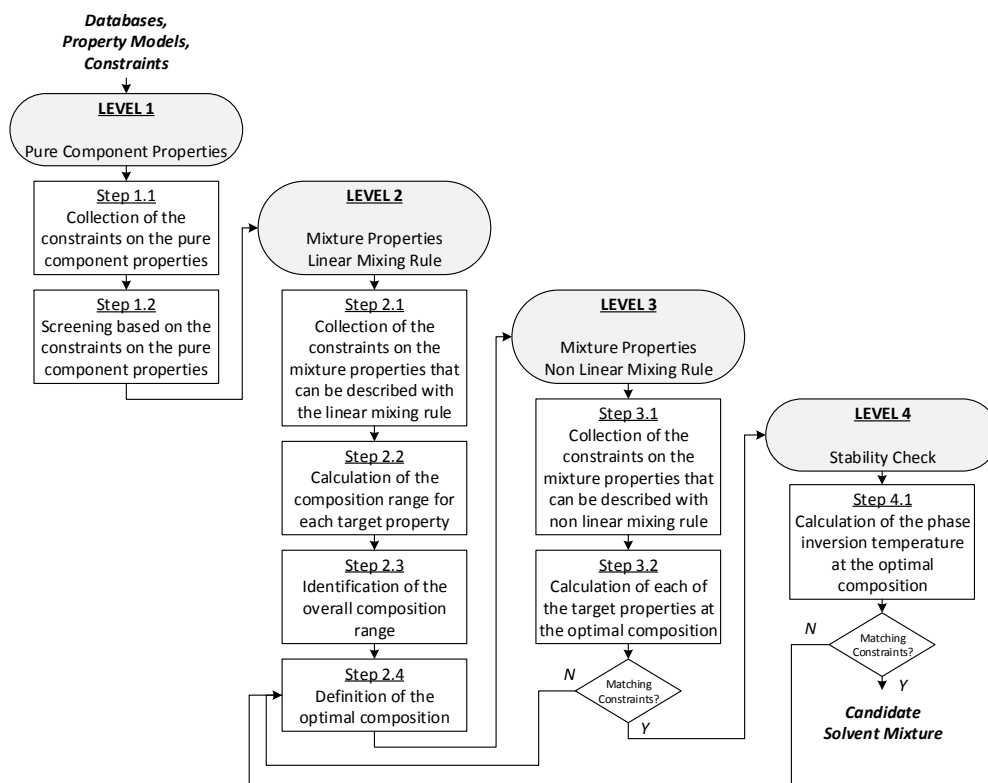
Figure 4.2 Flow-diagram of the EMUD algorithm,
developed for the design of emulsified solvent mixtures

Four subsequent levels are involved, related to the different target property constraints set in the problem definition: first, in level 1, the constraints on the pure component properties of the candidate ingredients are satisfied; then, in level 2, the constraints on the mixture properties that can be described with linear models lead to the identification of the feasible composition range for each candidate solvent mixture, and to the definition of an

optimal composition. Subsequently, in level 3, the satisfaction of the constraints on the mixture properties described with non-linear models at the optimal composition are checked, and finally, in level 4, the stability of the candidate solvent mixtures as emulsions is verified. Both at level 3 and at level 4, as shown in Figure 4.2, if the constraints are not matched, the relative candidate solvent mixtures are not simply discarded, but a new optimal composition is calculated instead.

For simplicity, the above mentioned algorithm is described here with both the aqueous solvent phase and the organic solvent phase containing only one component each (or a pseudo-component, such as an essential oil, for which average properties are available or can be calculated). However, it can easily be extended to multi-component solvent phases, by integrating in the structure of Figure 4.2 the algorithm for mixture design (MIXD) and mixture stability (STABILITY), as given by Conte et al. (2011). The integration of the above mentioned algorithms is briefly given in §4.3.2.

4.3.1 The EMUD Algorithm

In this paragraph, the algorithm for emulsified mixture design (EMUD) is presented in details. First, a clarification on the definition of the target properties constraints is given, then each algorithm level is presented in all its steps.

Target Property Constraints

The constraints on a generic target property ζ_i can be of three different types:

- Perfect match $\zeta_{i,PM}$ constraint:

$$\zeta_i = \zeta_{i,PM} \tag{4.5}$$

- Lower $\zeta_{i,L}$ or upper $\zeta_{i,U}$ boundary constraint:

$$\zeta_{i,L} < \zeta_i \tag{4.6}$$

$$\zeta_i < \zeta_{i,U} \tag{4.7}$$

- Lower and upper boundary constraints:

$$\zeta_{i,L} < \zeta_i < \zeta_{i,U} \tag{4.8}$$

All the cases above, however, can be reduced to the case of equation (4.8), by applying a slack of 0.5% on the value of $\zeta_{i,PM}$ of equation (4.5), and by setting an upper boundary to a very large positive number, and a lower boundary to a very large negative number for equation (4.6) and equation (4.7), respectively. This way, all the constraints on the target properties in this algorithm, consider both a lower and an upper boundary conditions.

Level 1: Pure Component Properties

The emulsified mixture design algorithm starts by screening the candidate ingredients on the basis of pure component properties. In fact, constraints are usually set on pure com-

ponent properties of both the candidate ingredients for the aqueous and the organic solvent phases, such as, the flash point, the toxicity parameter, and the solubility parameters; as well as on some pure component properties of the candidate surfactant, such as, the cloud point, the Krafft temperature, and the surface tension reduction.

STEP 1.1

In this first step, all the defined constraints on the target pure component properties are collected and distinguished between those relative to the aqueous solvent phase, to the organic solvent phase, and to the surfactant system, and the target property constraints are set as in equation (4.8).

In relation to the Hansen and Hildebrand solubility parameters, the assumption of Hancock et al. (1997), further applied in the field of chemical product design by Conte et al. (2010), is considered. According to the above mentioned assumption, compounds having similar solubility parameters are miscible with each other. The mathematical representation of the above mentioned criteria is given in equation (4.9) and (4.10), for the Hildebrand and Hansen solubility parameters, respectively.

- Hildebrand solubility parameter:

$$\delta_{T,S} - 3 < \delta_{T,i} < \delta_{T,S} + 3 \tag{4.9}$$

  Where $\delta_{T,i}$ is the target Hildebrand solubility parameter, and $\delta_{T,S}$ is the Hildebrand solubility parameter of the solute.

- Hansen solubility parameters:

$$\delta_{D,S} - 2 < \delta_{D,i} < \delta_{D,S} + 2 \tag{4.10}$$

  Where $\delta_{D,i}$ is the target Hansen dispersion solubility parameter, and $\delta_{D,S}$ is the Hansen dispersion solubility parameter of the solute. The same numerical constraints are set for the Hansen polar and hydrogen bonding solubility parameters.

STEP 1.2

Once all the constraints on the pure component properties have been collected, the candidate ingredients are screened accordingly from the relative databases. These, need to contain all the target properties collected in step 1.1; if these properties are not available, dedicated pure component property models are used, in order to fill the gap. If it is not possible to retrieve, either to calculate, all the necessary pure component properties, the candidate ingredient is excluded.

Level 2: Mixture Properties – Linear Mixing Rule

After a first screening of the basis of the pure component properties, in the second level of the algorithm, the target mixture properties are introduced. As highlighted in Chapter 3, the mixture properties can be estimated by means of linear models (linear mixing rule) or with more accurate, non-linear models. In level 2, the target mixture properties that can be described with linear mixing rules are considered.

STEP 2.1

First, the target mixture properties that are described with linear mixing rules are collected, together with the relative constraints, in terms of boundaries of acceptance. Often, the constraints on the molar volume (or density), the dielectric constant and the toxicity parameter are collected here.

STEP 2.2

In this second step, the composition boundaries relative to each of the target properties constraints collected at step 2.1 are calculated.

Under the assumption of pure component solvent phases, the emulsified solvent mixture consists of a minimum of 3 components: the aqueous solvent phase ingredient, the organic solvent phase ingredient and 1 or more surfactants. Therefore, a general target mixture property $\zeta_{i,m}$ for the emulsified solvent mixture can be expressed as from equation (4.11).

$$\zeta_{i,m} = \sum_{j=1}^{NC=3} x_j \cdot \zeta_{i,j} = x_A \cdot \zeta_{i,A} + x_O \cdot \zeta_{i,O} + x_S \cdot \zeta_{i,S} \qquad (4.11)$$

Where the subscript m stands for the mixture property, and the subscripts A, O and S stand for the aqueous solvent phase ingredient, the organic solvent phase ingredient and the surfactant, respectively.

However, since the surfactants are frequently the most expensive among the above mentioned ingredients, their concentration is usually kept as low as possible. As already underlined in 3.4.1, nevertheless, the minimum concentration of the surfactant, with respect to the content of the aqueous solvent phase, in order to form an emulsion is defined by the critical micelle concentration. A safety factor of 10 is usually employed in order to consider the uncertainty in the estimation of the critical micelle concentration, so that the surfactant concentration considered in this algorithm is given by equation (4.12).

$$x_S = 10 \cdot CMC_S \cdot x_A \qquad (4.12)$$

Therefore, for the sake of the calculations, the ternary system can be reduced to a binary system, where the concentration of the surfactant is kept constant with the concentration of the aqueous solvent phase, as illustrated in equation (4.13).

$$\zeta_{i,m} = x_A \cdot \left(\zeta_{i,A} + 10 \cdot CMC_S \cdot \zeta_{i,S}\right) + x_O \cdot \zeta_{i,O} = x_A \cdot \zeta_{i,A}^* + x_O \cdot \zeta_{i,O} \qquad (4.13)$$

Where $\zeta_{i,A}^*$ represents the contribution to the target mixture property $\zeta_{i,m}$ given by the aqueous solvent phase ingredient and the surfactant, the concentration of which is fixed to be ten times the critical micelle concentration.

Under the assumption above of pseudo-binary system, each target property constraint mathematically described as in equation (4.8) can be translated into a composition range, as given in equation (4.14). From now on, for simplicity, the algorithm is presented in relation to the aqueous solvent phase ingredient, but the same calculations can be applied to the organic solvent phase ingredient.

$$x_{i,L,A} < x_{i,A} < x_{i,U,A} \tag{4.14}$$

Where $x_{i,L,A}$ and $x_{i,U,A}$ are the lower and the upper boundaries for the composition of the aqueous solvent phase ingredient A, relative to the target mixture property i, respectively, and the numerical values are calculated as in equation (4.15) and (4-16).

$$x_{i,L,A} = \frac{\left(\zeta_{i,m,U} - \zeta_{i,O}\right)}{\left(\zeta_{i,A}{}^* - \zeta_{i,O}\right)} \tag{4.15}$$

$$x_{i,U,A} = \frac{\left(\zeta_{i,m,L} - \zeta_{i,O}\right)}{\left(\zeta_{i,A}{}^* - \zeta_{i,O}\right)} \tag{4.16}$$

The above mentioned method is applied for each of the target mixture properties collected at step 2.1.

STEP 2.3

As the output of the step 2.2, a list of lower and upper boundaries for the composition of the solvent mixtures are identified. In this step, the overall composition range is identified, so that all the constraints on the mixture properties described with the linear mixing rule are satisfied.

In order to do that, the strictest condition for the composition has to be identified, for each mixture. That is, the maximum value of the lower boundaries on the composition $x_{i,L,A}$ has to be selected, while the minimum among the upper boundaries on the composition $x_{i,U,A}$ is chosen, as illustrated in equation (4.17) and (4.18).

$$x_{L,A} = \max_i\left(x_{i,L,A}\right) \tag{4.15}$$

$$x_{U,A} = \min_i\left(x_{i,U,A}\right) \tag{4.16}$$

Where $x_{L,A}$ and $x_{U,A}$ are the overall lower and upper boundaries for the compositions of the aqueous solvent phase ingredient.

Depending on the numerical values of the two composition boundaries, three conditions can be encountered:

- Condition 1: $x_{L,A} < x_{U,A}$, then a feasible overall composition range exists, and it is described by equation (4.17).

$$x_{L,A} < x_A < x_{U,A} \tag{4.17}$$

- Condition 2: $x_{L,A} = x_{U,A}$, then the feasible overall composition range exists, and it is described by equation (4.18).

$$x_A = x_{L,A} = x_{U,A} \tag{4.18}$$

- Condition 3: $x_{L,A} > x_{U,A}$, then a feasible overall composition range cannot be identified, therefore the relative solvent mixture is rejected.

STEP 2.4

Given the overall composition range identified in step 2.3, in this step the optimal composition is chosen. In this algorithm, the optimal composition is defined as the overall composition value leading to the cheapest mixture, for all the candidate mixture that have not been rejected in step 2.3.

Under the assumption of pseudo-binary solvent mixture given in step 2.3, the solution of the problem is straightforward. In fact if the pseudo-component consisting of the aqueous solvent phase ingredient and the surfactant (at a concentration 10 times larger than its critical micelle concentration) is cheaper than the organic solvent phase ingredient, then the optimal composition is defined for $x_A = x_{U,A}$. If the organic solvent phase ingredient is the cheaper, instead, the optimal composition is defined for $x_A = x_{L,A}$.

Level 3: Mixture Properties – Non-Linear Mixing Rule

At level 3 of the EMUD algorithm the target mixture properties that cannot be described by means of the linear mixing rule are considered.

STEP 3.1

First, similarly to step 2.1, it is necessary to collect all the target mixture properties that are described with non-linear mixing rules, together with their constraints and boundaries of acceptance. The viscosity, the flash point and the surface tension are among the target mixture properties that are considered in this step. Phase equilibrium related properties, such as the fugacity, the activity, *etc.* are also, by definition, mixture properties that cannot be described with the linear mixing rule, but they are considered at the level 4 of the algorithm.

STEP 3.2

In this second step, non-linear constraints are applied to the different emulsified solvent mixtures at the optimal composition as defined in step 2.4, in order to further screening the candidates.

The target mixture properties are calculated at the given mixture composition, and at the temperature defined as the input to the whole algorithm. Then the numerical values are compared with the target values and the respective boundaries of acceptance.

If all the calculated target mixture properties match with the given constraints, then the candidate emulsified solvent mixtures are processed to level 4. If, on the contrary, one or more of the constraints on the target mixture properties is not matched, then the candidate solvent mixture at the optimal composition is rejected, and a new optimal composition is calculated in step 2.4, according to the following rule: "If the rejected emulsified solvent mixture had an optimal composition given by $x_A = x_{U,A}$, then the new optimal composition is defined by $x_A{}^{new} = x_{U,A} - 0.01$. If the rejected emulsified solvent mixture had the optimal composition given by $x_A = x_{L,A}$, then the new optimal composition is defined by $x_A{}^{new} = x_{L,A} + 0.01$.

Level 4: Stability Check

All the candidate emulsified solvent mixtures accepted to level 4 of the EMUD algorithm satisfy the constraints regarding all the target mixture properties. In level 4, the stability of the candidate emulsified solvent mixtures is verified.

STEP 4.1

In this only step of level 4, the stability of the candidate solvent mixture as emulsified products is verified by means of dedicated thermodynamic models, or correlations.

As the stability of an emulsion is a complex, multifaceted concept, in this level of the EMUD algorithm, the stability of the emulsified product as its ability to remain of the same type, that is, water-in-oil or oil-in-water, over a wide temperature range is investigated. In particular, the phase inversion temperature as the maximum (or minimum) temperature of stability is calculated.

If the calculated phase inversion temperature matches the relative constraint, then the candidate emulsified solvent mixture is verified to be matching all the given constraints, and it is therefore generated as final output of the algorithm. If, on the other hand, the calculated value for the phase inversion temperature does not match the given constraint, then a new optimal composition for the candidate emulsified solvent mixture is to be calculated in step 2.4, according to the rule defined in step 3.2.

4.3.2 Integration of the EMUD Algorithm with the MIXD Algorithm

As it has been highlighted in §4.3, the algorithm has been illustrated in §4.3.1 by assuming that both the aqueous and the organic solvent phase consist of a single component, or pseudo-component, each. Such an assumption, that allows an easier explanation of the whole EMUD algorithm, is however not very realistic, as often commercial products consist of complex mixture of aqueous and organic chemicals, respectively.

The EMUD algorithm, however, thanks to its modular structure, allows the integration with other algorithms, such as the MIXD algorithm and the STABILITY algorithms (Conte et al, 2010), specifically developed for the design of homogeneous solvent mixtures. It is not considered crucial to give a detailed description here of the integration of the different algorithms, while only the mutual interactions between the algorithms are illustrated and the integrated work-flow is shown.

In the stand-alone EMUD algorithm, the candidate aqueous and organic solvent phases, considered as mono-component, are screened in level 1. When the assumption of mono-component solvent phases is rejected, and the MIXD and STABILITY algorithms are added, they act by substituting the step 1.2 in relation to the aqueous and the organic solvent phases, while it remains as in §4.3.1 with respect to the surfactant system. The rest of the EMUD algorithm, then, remains as in Figure 4.2.

Figure 4.3 illustrates the integration of the different algorithms, by magnifying the overall work-flow at level 1 and 2.
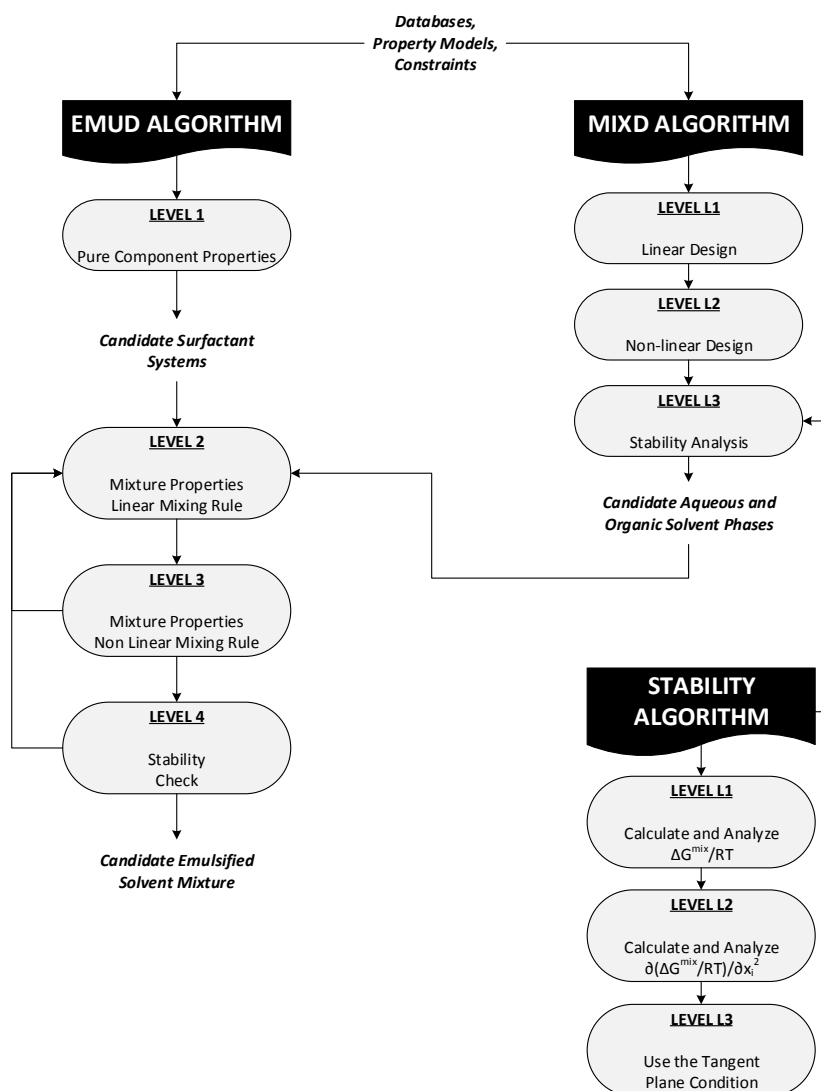
Figure 4.3 Flow-diagram of the integration between the EMUD, the MIXD and the STABILITY algorithms, for optimal design of emulsified solvent mixtures

4.4 Computer-Aided Tools

The solution of a chemical product design problem by applying systematic integrated model-based and experiment-based methodologies require the retrieval, the calculation, the use and the management of a wide range of models and procedure, as well as a very large amount of information and data. Computer-aided tools are necessary, in order to simplify the use of these procedure. In this paragraph, first the different toolboxes that have been used in this work, belonging to the Integrated Computer Aided System (ICAS) (Nielsen et al., 2001), are listed. Then, in §4.4.2, the Virtual Process-Product Design Laboratory (VPPD-Lab), and its use for the design of emulsified formulated products are briefly illustrated.

4.4.1 Miscellaneous ICAS Toolboxes

The ICAS software consists of a number of different toolboxes that help to efficiently solve a wide range of chemical engineering problems. In this work, in particular, the CAPEC DataBase, the Computer Aided Molecular Design tool (ProCAMD) and the Property Prediction tool (ProPred) have been extensively used.

- The CAPEC DataBase has been used as the basis for the creation of the ingredient databases (§4.1), and for the collection of experimental values relative to several target pure component properties;

- The ProCAMD has been used for the automatic generation of molecular structures to be added in the ingredient databases under development;

- The ProPred has been extensively used for the calculation of target pure component properties of several chemicals, when experimental values are not available. In ProPred, a large number of pure component property models and group contribution models are available, so that the needed properties can be predicted for a very wide range of chemicals. Almost all the pure component property models presented in §3.3 are available in ProPred and the addition of the pure component models dedicated to surfactant (§3.4) has been planned.

4.4.2 The VPPD-Lab for Emulsified Formulated Product Design

The Virtual Process-Product Design Laboratory has been originally developed by Conte et al. (2010) for the design/analysis of homogeneous formulated products and as a part of this work, it has been extended to emulsified formulated products. A unique feature of the framework is the use of problem specific templates, relative to the different type of products that can be designed/analyzed, through a generic design work-flow.

The VPPD-Lab allows to perform virtual experiments while searching for the most promising candidates. When these are found, the VPPD-Lab recommends experiments to verify the product formulations, thanks to the integrated knowledge-based (§4.2). That is, computer-aided techniques are used to search through a wide range of alternatives; just as a process simulator is able to simulate and analyze different chemical processes, the VPPD-Lab is able to design and analyze different chemical products.

The generic work-flow embedded in the VPPD-Lab, and illustrated in Figure 4.4, consists of a set of hierarchical steps and it is supported by a collection of tools, such as, a property model library (Chapter 3), a reliable knowledge-base (§4.1), structured databases (§4.2) and calculation routines (§4.3).

As from Figure 4.4, the workflow is based on a specifically developed ontology for knowledge representation covered by associated product attributes, their translation into properties, the corresponding property prediction models, and a wide range of data from different sources. In the first step, the product type is selected from a list of products available in the database: homogeneous formulations, gasoline blends, lubricant blends, jet fuel blends and emulsion-based formulations. In the second step, a set of product needs

is retrieved from the knowledge-base and augmented, if necessary. In the third step, the product needs are translated to product attributes (properties) by the knowledgebase, which also helps to define the property target values. The fourth step employs the problem specific formulation design in terms of selecting the ingredients (chemicals) and their amounts such that the property targets are satisfied. The ingredients are classified in terms of active ingredients (performs the main function or activity of the product), solvents (to dissolve the active ingredient and/or to deliver the product) and additives (to enhance the product quality). In the final fifth step, each feasible product formulation is verified through model-based tests to check for stability, performance enhancement, etc.

STEP 1

Select
Product Type

STEP 2

Identify
Consumer Needs

STEP 3

Translate
Product Needs

STEP 4

Select
Ingredients

STEP 5

Verify
Product
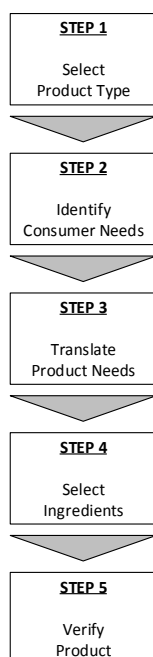
Figure 4.4 Generic work-flow
integrated in the VPPD-Lab

A large number of properties is needed, as well as a large number of property models to generate these values, when experimental data are missing- These are both implemented in VPPD-Lab, in terms of structured databases and property models, some of which are common to all the product types, while others are product specific. Properties like the cost, the density and the Hansen solubility parameters are necessary for all the different product design scenarios, and the property models employed are the same. Examples of product specific properties for the emulsified formulated products are the cloud point, the Krafft temperature and the critical micelle concentration.

The specific template for emulsified formulated product design follows a systematic procedure defined in terms of four tasks:

- Task 1, the problem definition, is responsible for the identification of the product needs, their translation into target properties and the specification of a set of appropriate target values for the identified target properties. These data are passed from steps 1-3 of the generic workflow.
- Task 2 selects the active ingredients of the formulation from the available database. The necessary property models and data are retrieved from the libraries and a rule-based procedure is employed to make the selections.
- Task 3 determines the solvent mixture. For emulsion-based products, a choice needs to be made of two solvent liquid phases (the aqueous solvent phase and the organic solvent phase). Also, a selection of an appropriate surfactant system, needed to keep the above mentioned liquids in the emulsified form, is made.
- Task 4 is responsible for the selection of additives. Finally the stability of the product is assessed by means of an in-house algorithm.

More details on the application of the VPPD-Lab to a conceptual case study for the design of an emulsified formulated product can be found in Appendix C.

5

FRAMEWORK
AND METHODOLOGY

Various models, as well as methods and tools are necessary for the design of emulsified formulated products, and these have been presented in Chapter 3 and 4, respectively. In this chapter, a systematic framework is presented, where the above mentioned models, methods and tools can operate in a synergic manner within the overall design work-flow and data-flow. The overall systematic methodology for the design of emulsified formulated products is therefore implemented within this framework, which has been further tested on several case studies, presented in Chapter 6.

The systematic methodology, integrating model-based as well as experiment-based techniques, requires the definition of the product of interest as the input, and then gives as the output a validated emulsified formulation, containing a list of ingredients together with their relative concentrations.

In §5.1, an overview of the framework is given, with particular attention to the definition of the work-flow and on the interactions between the different design stages. The problem definition stage is presented in details in §5.2, followed by the model-based stage for product synthesis and design in §5.3 and by the experiment-based stage for verification and further refinement in §5.4.

5.1 The Framework

In this work an integrated methodology for emulsified formulated product design, consisting of a model-based stage and an experiment-based stage interacting with each other, is presented. The methodology is inspired by the approach proposed by Ng, Gani and

Dam-Johansen (2007), further extended for the solution of problems involving homogeneous formulated products by Conte et al. (2010).

The methodology consists of three stages, as shown in Figure 5.1, where the necessary methods and tools differ for each of the stages. They are briefly presented here, while a better description of each is given in the next paragraphs.

INPUT
Type of Product

STAGE 1
Problem Definition Stage

Task 1.1 Consumer Needs
Task 1.2 Target Properties I
Task 1.3 Target Properties II

STAGE 2
Model-based Stage

Task 2.1 Active Ingredients
Task 2.2 Emulsified Solvent Mixture
Task 2.3 Additives
Task 2.4 Product Properties and Stability

STAGE 3
Experiment-based Stage

Task 3.1 Design of Experiments
Task 3.2 List of Actions
Task 3.3 Experimental Work
Task 3.4 Validation

OUTPUT
Validated Formulation

Figure 5.1 The work-flow of the overall integrated methodology
for the design of emulsified formulated products

The problem definition stage is very important, as it is first in the hierarchical structure of the methodology, and backwards interactions from the other stages are not possible. That is, any decision taken at this level influences the decisions taken in the following steps, but not vice versa. The main tool implemented in this stage is the knowledge base, presented in Chapter 4. At this point, a list of target thermo-physical properties (including target values and boundaries of acceptance) and a list of necessary categories of ingredient are generated, given the type of product of interest as the input.

The model-based stage takes the results of the problem definition stage and, through the use of property models, structured databases and dedicated algorithms, converts them into

a candidate emulsified formulated product to be verified in the next stage, that is, a list of candidate ingredients and relative concentrations are determined. In order to reliably screen thousands of candidates while applying the "reverse design" technique, the model-based stage employs a decomposition strategy, so that the solution method is divided into a set of sub-problem to be solved individually.

Finally, the experiment-based stage is in charge of verifying the formulated product proposed by the previous stage, through focused experiments. A list of actions is also generated so that, if the results of these experiments do not match with the expected results provided by the previous stage, appropriate corrections can be taken.

As illustrated in Figure 5.1, the interaction between the model-based stage and the experiment-based stage is not unilateral, and iterations within the two stages are required if the formulation proposed by the model-based stage is not validated in the experiment-based stage. Iterations are necessary until a candidate formulated product generated by the model-based stage is verified by the experiment-based stage.

The main objective of the overall integrated methodology is to quickly and efficiently screen between many candidates by means of mathematical models and algorithms, so that the valuable experimental resources are reserved for the final verification and possible refinement. This, however, is possible only if a set of product needs, in terms of target properties to be satisfied by and necessary categories of ingredients to be included in the candidate emulsified formulated product, is generated.

5.2 Stage 1: Problem Definition Stage

The first stage of the integrated methodology for the design of emulsified formulated products consists of problem definition. Based on the knowledge base presented in Chapter 4, this stage defines a set of target properties and ingredients relative to the product to be designed. It involves three main tasks, containing different sub-tasks. Figure 5.2 shows the detailed work-flow of the problem definition stage.

5.2.1 Task 1.1: Consumer Needs

> _Input:_ Type of product.
> _Tools:_ Knowledge base.
> _Output:_ Main and secondary consumer needs.

In this first task, the consumer needs (ψ_i) relative to the product to be designed are collected from the knowledge base and they are divided into main (ψ_M) and secondary (ψ_S) needs.

Subtask 1.1.1: Collection of the Consumer Needs

The consumer needs are defined as all those attributes that consumers need and/or want from a product in order to buy it. The consumer needs include a wide variety of product

attributes, ranging from cosmetic properties and sensorial factors, to safety and toxicity related property, to product form and phase stability.

INPUT
Type of Product

TASK 1.1
Consumer Needs

Sub-task 1.1.1 Collection of the Consumer Needs
Sub-task 1.1.2 Definition of the Main and Secondary Consumer Needs

TASK 1.2
Target Properties I

Sub-task 1.2.1 Collection of the Target Properties I – Not Modelled
Sub-task 1.2.2 Definition of the Active Ingredients and of the Additives

TASK 1.3
Target Properties II

Sub-task 1.3.1 Collection of the Target Properties II - Modelled
Sub-task 1.3.2 Classification of the Target Properties
Sub-task 1.3.3 Identification of the Target Values and of the Boundaries of Acceptance

OUTPUT
Consumer needs, categories of active ingredients and additives, target properties with target values and boundaries of acceptance

Figure 5.2 The work-flow of the problem definition stage
for the design of emulsified formulated products

This subtask is of fundamental importance in the industry, as usually the difference between an adequate and a wrong choice of the consumer needs to be satisfied is very subtle, but it may lead to the synthesis and marketing of a blockbuster versus an unsuccessful product. Therefore, in the industry, it is common practice to invest a lot of time and resources for product conceptualization, by means of market surveys, analysis of the competing products and so on (Cheng et al., 2009). The management and sales and marketing units are usually responsible for such activities, as the required skills are different than those characterizing the formulation chemists/engineers.

Therefore, in this work, the data and information relative to the collection of the consumer needs for a range of formulated products have been retrieved from literature, patents, real products, insight, industrial experience and common sense. They have been then collected and stored in the knowledge base (§4.2), so that they can be provided, when needed in an efficient way.

Subtask 1.1.2: Definition of the Main and the Secondary Consumer Needs

The main needs (ψ_m) are defined as the main reasons for which consumers would buy the product. For example, the main need of an insect repellent product is to repel mosquitos, while the main need of a UV sunscreen, which has been used as a conceptual case study for this methodology and it is presented in Chapter 6, is to provide protection from the UV radiations.

The secondary needs (ψ_s), on the other hand, are defined as accessory product attributes which can drive consumers to prefer a product, as compared to another, but the absence of which does not influence the minimum product performances. These secondary needs are usually related to the form of the formulated product (liquid, solid, emulsion, spray…), safety, toxicity, cosmetic properties (odor, color…), and so on. For example, between the secondary needs of an insect repellent product, one can identify a pleasant odor and the ease of spreading or spraying, while in relation to a UV sunscreen, a pleasant skin feeling and the prevention of skin ageing are among the secondary needs.

5.2.2 Task 1.2: Target Properties I

> *Input:* Main and secondary consumer needs.
> *Tools:* Knowledge base.
> *Output:* Categories of active ingredients and additives.

In this second task, first those consumer needs which cannot be translated into a set of thermo-physical properties are identified. These needs, therefore, are translated into properties that directly link to categories of ingredients (ξ_i) that are known to be able to satisfy these needs. Then, the distinction between main and secondary consumer needs is used to distinguish between the active ingredients (ξ_{AI}) and the additives (ξ_{ADD}).

Sub-task 1.2.1: Collection of the Target Properties I – Not Modelled

There are several product attributes, in fact, that are very difficult to be quantitatively described by means of thermo-physical properties. A simple example regards color and odor, but also for important product specific needs, such as the ability of repel mosquitos for insect repellent lotions, for instance, there are no models able to describe such attributes.

However, there are categories of ingredients that have been recognized to be able to satisfy these needs, such as coloring agents (dyes or pigments), in relation to the main need of a paint. There are cases, however, where more classes of ingredients correspond to a

single consumer need, such as in relation to the protection from UV radiations. In fact, there are very few chemicals which have been proven to be effective in screening the skin from both UV-A and UV-B radiations, therefore UV-A as well as UV-B absorbers and/or filters are necessary.

In the design methodology, then, these consumer needs are translated into target properties, even though no quantitative models are available. These properties are defined in this work as target properties I, in order to distinguish them from those properties (target properties II) for which models are available. For the application of such a sub-task, a reliable knowledge base is necessary.

Sub-task 1.2.2: Definition of the Active Ingredients and of the Additives

The distinction between main and secondary needs defined in sub-task 1.1.2 is considered in this sub-task, together with the target properties collected in sub-task 1.2.1, in order to distinguish between active ingredients and additives.

The active ingredients (ξ_{AI}) are defined as those ingredients that are able to satisfy the properties connected to the main product needs (ψ_m), while the additives (ξ_{ADD}) are defined as those ingredients that are able to satisfy the target properties connected to the secondary product needs (ψ_s). Such a distinction is of fundamental importance as, in the model-based stage, the active ingredients and the additives are considered in two different steps of the framework, with one (the selection of the active ingredients) indirectly influencing the other (the selection of the additives). For the application of this sub-task, structured databases are necessary, together with a reliable knowledge-base, linking the target properties to the relative classes of ingredients.

5.2.3 Task 1.3: Target Properties II

> *Input:* Consumer needs, categories of active ingredients and additives.
> *Tools:* Knowledge base.
> *Output:* Target properties, target values, boundaries of acceptance.

In this third task, those consumer needs which can be translated into a set of thermo-physical properties (ζ_i), that is, target properties are identified. Then, they are divided between those relative to the active ingredients (ζ_{AI}), those relative to the additives (ζ_{ADD}), those relative to the final product (ζ_{PROD}), and those relative to all the ingredients (ζ_{ALL}). Finally, target values and boundaries of acceptance are set for each target property.

Sub-task 1.3.1: Collection of the Target Properties II – Modelled

The range of product attributes for which a direct translation into one or a combination of thermo-physical properties is possible, is identified here. Examples are those consumer needs related to toxicity, where properties as the lethal concentration are usually employed, or to the type of product, where properties as the vapor pressure of the melting temperature are used to make sure the product is in the desired physical form.

In this sub-task, the consumer needs are translated, through the knowledge base, into thermo-physical properties (ζ_i), defined as target properties II, for which quantitative models are available.

Sub-task 1.3.2: Classification of the Target Properties

The target properties collected in the previous sub-task are to be classified into four different classes:

- ζ_1 consist of those properties relative to the active ingredients, such as the surface tension and the critical micelle concentration in relation to the foam-ability of a hand-wash (§6.3);
- ζ_2 consist of those properties relative to the additives, referring to specific ingredient performances such as the solubility in the solvent mixture, quantified by means of the solubility parameters;
- ζ_3 consist of those properties relative to all the ingredients of the formulations; examples are the flash point and the toxicity parameter, that are fundamental properties, in relation to safety and toxicity issues;
- ζ_4 consist of those properties relative to the overall product; they are mainly bulk mixture properties such as density, viscosity and surface tension, derived from consumer needs like spread-ability, spray-ability, *etc.*

Sub-task 1.3.3: Identification of the Target Values and of the Boundaries of Acceptance

For the set of target properties ζ_i identified in sub-task 1.3.1 and further classified in sub-task 1.3.2, the target value ($\zeta_{i,T}$) and the boundaries of acceptance($\zeta_{i,L}$, $\zeta_{i,U}$) are defined.

The target values are the numerical values that certain properties are desired to match, while the boundaries of acceptance are the lower and upper numerical values that certain properties are required not to go beyond. The boundaries of acceptance are usually applied for screening purposes, while the target values are commonly used as performance index for ingredient selection.

It has to be underlined that target values are not always required, while two boundaries of acceptance are necessary for each target property. In case only a lower or an upper boundary can be set, the second is set to a very large positive value, or to a very negative large value, respectively.

5.3 Stage 2: Model-based Stage

The model-based stage is the second stage of the integrated methodology for the design of emulsified formulated products. The property models and algorithms presented in the previous chapters are used here, so that the target properties and ingredients generated in the previous stage are translated into a candidate formulation. Similarly to the previous, this is also a step-by-step stage, and it consists of four main tasks, each containing different sub-tasks. Figure 5.3 illustrated in details the relative work-flow.

INPUT
Categories of active ingredients and additives, target properties with target values and boundaries of acceptance

TASK 2.1
Active Ingredients

Sub-task 2.1.1 Screening of candidate Active Ingredients
Sub-task 2.1.2 Selection of the best Active Ingredients
Sub-task 2.1.3 Collection of the Active Ingredients properties

TASK 2.2
Emulsified Solvent Mixture

Sub-task 2.2.1 Retrieval of the Databases
Sub-task 2.2.2 Choice of the Property Models
Sub-task 2.2.3 Application of the EMUD Algorithm
Sub-task 2.2.4 Collection of the Emulsified Solvent Mixture properties

TASK 2.3
Additives

Sub-task 2.3.1 Screening of the Additives
Sub-task 2.3.2 Selection of the best Additives
Sub-task 2.3.3 Collection of the Additives Properties

TASK 2.4
Product Properties and Stability

Sub-task 2.4.1 Calculation of the Product properties
Sub-task 2.4.2 Evaluation of the Product stability

OUTPUT
Candidate Product with thermo-physical properties

Figure 5.3 The work-flow of the model-based stage
for the design of emulsified formulated products

5.3.1 Task 2.1: Active Ingredients Selection

Input: Categories of active ingredients, target properties of active ingredients.

Tools: Databases, property models.
Output: Candidate active ingredients with thermo-physical properties.

The objective of this task is to select the most advantageous active ingredients (ξ_{AI}). First, the appropriate databases are retrieved, according to what has been defined in task 1.2; then, the selection of the optimal ingredients is performed on the basis of their target properties (ζ_1) and the relative target values, defined in task 1.3; and finally, all the relevant thermo-physical properties of the candidate ingredients, if not available in the database, are calculated through dedicated property models.

Sub-task 2.1.1: Screening of the Candidate Active Ingredients

In this sub-task, the relative databases to the necessary categories of active ingredients (ξ_{AI}) are retrieved from the database library or, if not available, are generated with information from literature, patents, real products and experience.

The database needs to be filled with all the relevant properties for the selection task, that is, at least all the target properties relative to the active ingredients (ζ_1) and to all the ingredients (ζ_3), as they have been defined in sub-task 1.3.2. If these properties are not available, property models are applied to fill the gap in the property table.

The screening process is carried out by screening out all the ingredients of the database whose numerical value of any of the target properties is lower than the lower boundary of acceptance ($\zeta_{i,L}$), or higher than the upper boundary of acceptance ($\zeta_{i,U}$). Those ingredients for which it has not been possible to retrieve either to calculate all the necessary target properties are also excluded from the selection process.

Sub-task 2.1.2: Selection of the Best Active Ingredients

From the short-list of candidate active ingredients generated in sub-task 2.1.1, this sub-task is in charge of selecting the most appropriate ingredient to be added to the formulation, based on a performance index. The performance index is defined on the basis of one or more of the following criteria:

- Effectiveness: the most effective chemical is selected;
- Safety and health: the safest chemical is selected;
- Environment: the most environmental friendly chemical is selected;
- Cost: the cheapest chemical is selected;
- A combination of the above criteria.

The target properties usually employed in this sub-task are ζ_1 for the effectiveness, the flash-point for the safety, the toxicity parameter for the environment, and the market price for the cost.

Sub-task 2.1.3: Collection of the Active Ingredients Properties

For each of the active ingredients selected in sub-task 2.1.2, it is required in this sub-task to retrieve from the relative database all the properties that will be necessary in the rest

of the model-based stage. If these are not available in the database, property models are used for estimation.

In addition to the target properties needed in sub-task 2.1.2, in order to perform the selection, here it is necessary to collect information about the solubility of each active ingredient (usually quantified by means of the Hildebrand and/or Hansen solubility parameters), density, viscosity, and any other thermo-physical properties that can influence the bulk properties ζ_4.

Finally, as most often models able to relate the active ingredients performances (ζ_1) to their concentration in the solvent mixture are not available, it is necessary to retrieve from the knowledge base the recommended concentration of each of the active ingredients selected. Such concentration are eventually changed in the experiment-based stage, in order to find the optimal solution, which corresponds to the minimum amount of the active ingredients needed, in order to satisfy the wanted performances (ζ_1).

5.3.2 Task 2.2: Emulsified Solvent Mixture Design

> _Input:_ Target properties of the product.
> _Tools:_ Database. Property models, EMUD algorithm.
> _Output:_ Candidate emulsified solvent mixture, with thermo-physical properties.

In this task, the optimal solvent mixture in the emulsified form is designed, that is, the most advantageous aqueous solvents, organic solvents, surfactant system and their composition. First, the relative structured databases are selected, then the necessary property models are chosen from the property model library, and finally the emulsified solvent mixture design (EMUD) algorithm is applied, together with the information generated in sub-task 1.3.2 relative to the overall product (ζ_4) and to all the ingredients (ζ_3).

Sub-task 2.2.1: Retrieval of the Databases

The categories of ingredients to be included in an emulsified solvent mixture are defined _a priori_, rather than in sub-task 1.2.1 as for the other ingredients of the formulation. Such a solvent mixture, in fact, requires an aqueous solvent (or solvent mixture), an organic solvent (or solvent mixture) and a surfactant system.

In this sub-task, the relative databases are retrieved, or generated if not available. Similarly to sub-task 2.1.1, the databases need to be filled with all the relevant properties: ζ_4 and ζ_3; if they are not, dedicated property models are applied to fill all the gaps. Every chemical in the database for which any of the above mentioned target properties is missing, is screened out from the design process.

Sub-task 2.2.2: Choice of the Property Models

In a mixture design calculation, pure component property models and mixture property models are simultaneously applied to solve the reverse problem, that is, given a set of

constraints ($\zeta_{i,T}$, $\zeta_{i,L}$, $\zeta_{i,U}$) on a set of target properties (ζ_4, ζ_3), determine the solvent mixture that match the constraints. However, as described in Chapter 3, there are some target properties for which different property models can be applied. This is often true, for example, for mixture properties, where linear mixing rule or rigorous models can be applied. In this sub-task, it is necessary to choose which property models are to be applied by the EMUD algorithm for the solution of the reverse problem. It is common practice, however, to use linear mixing rule models in the algorithm since the rigorous models are computationally expensive and it is tedious to apply them to several candidate solvent mixtures. On the contrary, in the verification step, where the thermo-physical properties of the chosen solvent mixture are calculated, rigorous models are applied.

Sub-task 2.2.3: Application of the EMUD Algorithm

In this sub-task, the emulsified solvent mixture design (EMUD) algorithm is applied. The database to be used are defined in sub-task 2.2.1, while the relevant target properties, together with the relative target values and boundaries of acceptance, defined in sub-task 1.3.3, are collected in sub-task 2.2.2.

It has to be underlined that applying the constraints on the target properties relative to the product (ζ_4) to the solvent mixture design, corresponds to assuming that the bulk properties of the product are identified with those of the solvent mixture. Such an assumption is necessary, in order to reduce the complexity of the problem, but for most of the consumer oriented products, the solvent mixture represents 70-90% (in mole percentage) of the whole product, therefore this assumption is very close to reality.

The output of this sub-task is a candidate emulsified solvent mixture, with its optimal composition, as defined by the EMUD algorithm (§4.3).

Sub-task 2.2.4: Collection of the Solvent Mixture Properties

In this sub-task, rigorous mixture property models are applied to the candidate emulsified solvent mixture generated in sub-task 2.2.4, in order to calculate and collect all its relevant properties. In this sub-task, it is possible to compare the results obtained with the application of the linear mixing rule to a certain set of target properties with those obtained with the rigorous models. More important, here it is necessary to check that, by applying the rigorous models, the target properties (ζ_4, ζ_3) are still matching the given constraints ($\zeta_{i,T}$, $\zeta_{i,L}$, $\zeta_{i,U}$).

If this is not true, then the candidate solvent mixture is rejected, and the sub-task 2.2.3 is run again, without the solvent mixture which has just been excluded.

5.3.3 Task 2.3: Additives Selection

Input: Categories of additives, target properties of additives, candidate emulsified solvent mixture.

Tools: Databases, property models.

Output: Candidate additives with thermo-physical properties.

In this third task, the best chemicals for satisfying the secondary consumer needs (ψ_s) are selected. Similarly to task 2.1, first the necessary databases relative to the categories of ingredients identified with ξ_2, are retrieved. Then, the most advantageous chemicals for each category of ingredients are selected by means of an appropriate performance index and finally, all the relevant thermo-physical properties of the candidate additives are retrieved from the databases, or they are calculated through dedicated property models, if not available.

Sub-task 2.3.1: Screening of the Additives

The necessary categories of ingredients satisfying the secondary needs (ψ_S), that is, the additives (ξ_2) have been identified in sub-task 1.2.2. Here, the adequate databases are retrieved or generated, if any of them is not available in the database library.

These databases need to contain all the relevant properties for the selection task, that is: the target properties relative to the additives (ζ_2), those relative to all the ingredients (ζ_3), and solubility information, usually quantified by Hansen and Hildebrand solubility parameters. Differently than with active ingredients, in fact, the additives are selected after the solvent mixture, therefore proper constraints on the solubility parameter (see §4.3) are needed. If any of these information is not available, property models have to be applied.

In this sub-task, then, a first screening among the candidate additives is performed. All the chemicals failing to match all the constraints given by the lower ($\zeta_{i,L}$) and the upper ($\zeta_{i,U}$) boundaries of acceptance, are in fact discarded. Moreover, any ingredient missing a numerical value for any of the relevant target properties is also excluded from the selection process.

It is easy to imagine that, in this sub-task, most of the discarded ingredients are related to the constraints on the solubility parameters. This, however, matches well the reason behind the generation of the step-wise methodology. In fact, the active ingredients are recognized as the most important ingredients of the formulation, as they satisfy the main consumer needs. Therefore, their selection must be performed without further constraints. Consequently, the solvent mixture is designed to be able to efficiently dissolve the active ingredients, besides matching the constraints on the target properties and in a way, then, the design of the solvent mixture is "active ingredients oriented"; so, the possible exclusion of some additives because of solubility issues in the designed solvent mixture must be regarded as an indirect consequence of the selection of the active ingredients.

This fact underlines the importance of the problem definition stage, and in particular of the definition of the main and secondary consumer needs (sub-task 1.1.2).

Sub-task 2.3.2: Selection of the Best Additives

In this sub-task, the most advantageous additives are selected from the short-list of candidates generated in sub-task 2.3.1. The selection is carried out with the same criteria

illustrated in sub-task 2.1.1, that is, based on a performance index defined on the basis of one or more of the following criteria:

- Effectiveness: the most effective chemical is selected;
- Safety and health: the safest chemical is selected;
- Environment: the most environmental friendly chemical is selected;
- Cost: the cheapest chemical is selected;
- A combination of the above criteria.

The target properties usually employed in this sub-task are ζ_2 for the effectiveness, the flash-point for the safety, the toxicity parameter for the environment, and the market price for the cost.

Sub-task 2.3.3: Collection of the Additives Properties

This sub-task is in charge of retrieving from the databases all the thermo-physical properties that are required in the next steps of the model-bases stage. These are, besides the target properties used in sub-task 2.3.1 and 2.3.2, density, viscosity, surface tension, *etc.*, as they influence the overall product properties ζ_4.

Even for additives, there are very few cases in which it is possible to relate the performances of the different ingredients ζ_2 with their concentration in the product. Therefore, it is necessary in this sub-task to define a recommended concentration, that can be eventually further optimized in the experiment-base stage.

5.3.4 Task 2.4: Product Properties and Stability Check

> *Input:* Candidate active ingredients, emulsified solvent mixture, and additives.
> *Tools:* Property models.
> *Output:* Candidate emulsified formulated product with thermo-physical properties.

In this fourth task, a model-based stability check on the overall candidate formulation is performed, and the relevant thermo-physical properties of the overall product are calculated and compared with the requirements: ζ_{PROD}. In this task, differently than in sub-task 2.2.4, all the formulation ingredients are included in the mixture property calculations. In fact, the addition to the formulation of certain categories of additives can dramatically change both the overall product properties, and the product stability.

Sub-task 2.4.1: Calculation of the Product Properties

The additives are necessary in the formulation, so that some of the secondary consumer needs are satisfied. There are categories of additives that are responsible of remarkable effects on some overall thermo-physical product properties, even if added in very low concentrations. Therefore, it is necessary to check, once all the ingredients of the formulation have been chosen, if the overall product target properties (ζ_4) are still in agreement

93

with the boundaries of acceptance set in sub-task 1.3.3. That is, if the assumption that the solvent mixture is responsible for the overall product properties is verified.

A common example of such a category of additives are the thickeners. Thickeners are substances which increase the viscosity of a liquid mixture without substantially modifying any other properties. They usually consists of synthetic as well as natural polymers, and they are used in the range of 0.2-2% in mole percentage.

In this sub-task, all the necessary product target properties (ζ_4) are calculated using rigorous mixture property models and, if the calculated properties are found to be out of the boundaries of acceptance, the candidate formulation is rejected. It is then necessary to identify the reason for this discrepancy and find the source. Corrective actions are taken going backward in the model-based stage, that is, first applying again task 2.3. If no possible correction can alter the product target properties enough to match the given boundaries of acceptance, then task 2.2 is applied again, and the rigorous models need to be used in the EMUD algorithm.

Sub-task 2.4.2: Evaluation of the Product Stability

Some categories of additives contain, among all the candidates, polar and/or ionic species. These chemicals, particularly if the designed surfactant system contains one or more ionic surfactant, can be responsible of extreme changes in the stability of the product as an emulsion.

Therefore, in this sub-task, the product stability as an emulsion is checked. Compared to the stability check performed in the EMUD algorithm (§4.3), here all the ingredients are considered, while the model applied is the same (§3.3.2).

If the product stability is not confirmed, corrective actions are taking by changing the selection of the additives, as those ingredients that are considered responsible of the product instability are rejected in sub-task 2.3.1. If the product stability is confirmed, then a candidate emulsified formulated product has been generated by the model-based stage, and it needs to be validated in the experiment-based stage.

5.4 Stage 3: Experiment-based Stage

The third stage of the integrated methodology for the design of emulsified formulated products is the experiment-based stage. Here, targeted experiments are performed in order to validate the candidate formulation generated in the model-based stage. If the product is not validated, however, this stage is in charge of suggesting refinements to the proposed formulation, by a proper list of action. This stage consists of four tasks, and several sub-task, as illustrated in Figure 5.4.
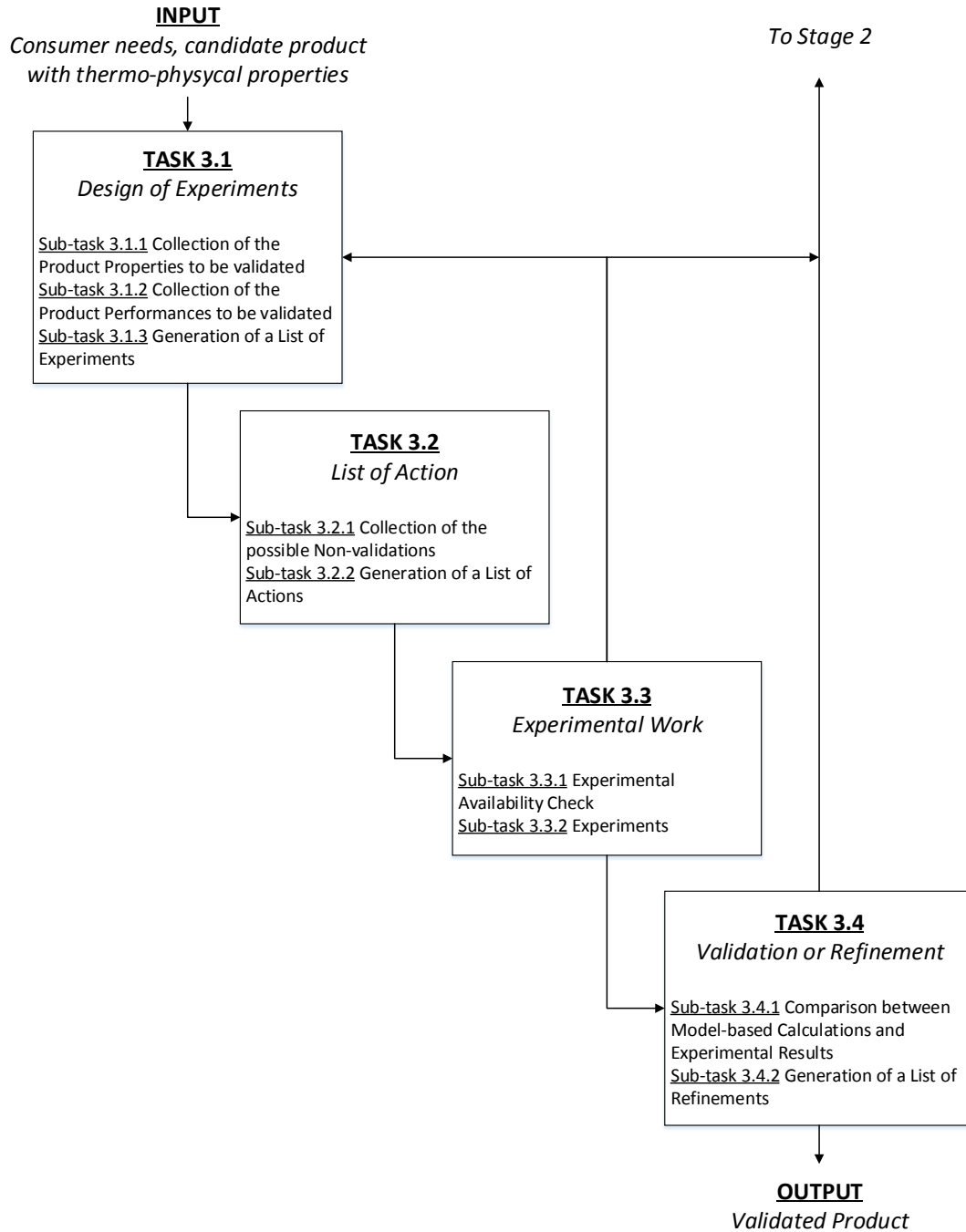
Figure 5.4 The work-flow of the experiment-based stage
for the design of emulsified formulated products

5.4.1 Task 3.1: Design of Experiments

Input: Consumer needs, candidate emulsified formulated product.
Tools: Knowledge base.
Output: List of experiments.

In this first task of the experiment-based stage, a list of experiments to be performed is generated, given the consumer needs defined in sub-task 1.1.1 (ψ_i) and their translation into target properties (ζ_i) and necessary classes of ingredients (ξ_i). Depending on the above mentioned distinction, experiments for product properties validation and experiments for product performances validation are defined, respectively.

Sub-task 3.1.1 Collection of the Product Properties to be validated

As it has been underline in the problem definition stage, some of the consumer needs collected in task 1.1 can be translated into target properties, that is, into a set of thermo-physical properties with appropriate target values and/or boundaries of acceptance. All the thermo-physical properties involved in the mode-based stage (ζ_i) are collected here, as experiments are necessary to validate the model-based predictions.

Sub-task 3.1.2 Collection of the Product Performances to be validated

In relation to some of the consumer needs collected in task 1.1, however, it is not possible to find correspondent thermo-physical properties able to numerical describe their satisfaction, or there are no models available able to quantify such a property. Therefore, such consumer needs have been translated, thanks to the knowledge base, into categories of ingredients that are known to be able to provide these product performances. All the consumer needs for which a direct translation into thermo-physical target properties has not been possible, therefore, are collected in this sub-task.

Sub-task 3.1.3 Generation of a List of Experiments

In this sub-task, the product properties (collected in sub-task 3.1.1) and the product performances to be validated (collected in sub-task 3.1.2) are translated into a set of experiments to be performed, and therefore a list of experiments for the product validation is generated.

In relation to the product properties, most of the experiments consists in measurements of bulk product properties (such as density, surface tension, viscosity, *etc.*), of properties related to the phase stability (solubility, mutual miscibility, *etc.*) and, when possible, of specific properties related to the product itself, such as the sun protection factor (SPF) for a UV sunscreen.

In relation to the product performances, instead, the validation experiments consists of a set of panel test, where product attributes such as the color, the odor and the feeling on the skin are validated by direct observation. Usually, among these experiments, also properties which have not been numerically modeled are verified, such as the pH, and the product shelf-life is also obtained by means of accelerated experiments.

The above mentioned experiments are arranged so that those experiments which are simple and non time consuming are performed first, so that if problems are identified during the experimental work, the procedure can stop and proper amendments are taken before the difficult tests, requiring big employment of time and/or resources are performed.

5.4.2 Task 3.2: List of Actions

Input: Candidate emulsified formulated product, list of validation experiments.
Tools: Knowledge base.
Output: List of actions.

In this second task, a list of action is generated in order to refine the candidate formulated product in case the experiments defined at task 3.1 do not validate it. Each experiments is then coupled with an action to be performed in case the results of the experiments are not satisfactory.

Sub-task 3.2.1 Collection of the possible Non-validations

Each of the experiments defined in sub-task 3.1.3 can provide results that do not validate the candidate formulated product generated by the model-based stage. In this sub-task, it is necessary to define and collect all the possible reason for non-validations, such as, for example, property measurements below the lower boundary or above the upper boundary of acceptance, or panel test having negative results.
Each of the experiments of the list generate in sub-task 3.1.3, therefore, must be accompanied by one or a list of reasons for non-validation.

Sub-task 3.2.2 Generation of a List of Actions

In this sub-task, the list of possible non-validations generated in sub-task 3.2.1, is coupled it with a proper list of actions. This list of actions consists of a set of amendments that are recommended, in order to refine the products, in case of non-validations. There may be more than one action, coupled to a single non-validation, as well as the same action can be coupled to more than one reasons for non-validation.
For example, the addition of a proper thickener is an action usually suggested when the measured viscosity is lower than the lower boundary of acceptance. However, for an emulsified formulated product, also a little increase of the amount on the dispersed phase in the emulsified solvent mixture, is recommended to increase the overall viscosity of the product. At the same time, the addition of some particular thickeners is reported to have sensible effects on the product stability, while the modification of the water-oil ratio of the emulsified solvent mixture influences all the bulk properties of the product, such as density, surface tension, *etc.*
The generation of a reliable list of action is possible thanks to thorough literature reviews, experimental evidences, industrial experience and common sense, all collected and stored in the knowledge base, which has been presented in §4.2.

5.4.3 Task 3.3: Experimental Work

Input: List of experiments.
Tools: Experimental set-up.
Output: Experimental results.

In this task, the experiments defined in task 3.1 are performed, subject to the availability of the necessary chemicals and experimental set-up.

Sub-task 3.3.1 Experimental Availability Check

In this first sub-task, the availability of the chemicals is checked, and the experimental set-up is verified. Therefore, proper decisions about the detailed experiments to be performed are taken.

If some of the necessary chemicals are not available, it is possible to go back to the model-based stage, to choose a proper replacement. If, on the contrary, it is not possible to perform one or more experiments because the relative experimental set-up cannot be used, then the work-flow goes back to Task 3.1, so that an alternative list of experiments can be generated.

This sub-task, then, produces a final list of experiments, by refining, if needed, the original list of experiments generated in sub-task 3.1.3.

Sub-task 3.3.2 Experiments

In this sub-task, the experiments collected in the final list of experiments are performed. As it has been underlined in sub-task 3.1.3, the experiments are arranged in a way that the least time and resource consuming experiments are performed first, so that, in case of discrepancies, it is not necessary to do those experiments that, on the contrary, are tedious and resource demanding.

All the results of the experiments are then collected together.

5.4.4 Task 3.4: Validation or Refinement

> _Input:_ Consumer needs, target properties, experimental results, list of actions.
> _Tools:_ -
> _Output:_ Validated emulsified formulated product or corrections to the model-based stage

In this last task, the results of the experiments performed in task 3.3 are compared with the expected results defined in the model-based stage. If all the experimental results are satisfactory, then the candidate formulated product generated through the model-based stage is validated. On the contrary, if any of the experiment is not satisfactory, the list of action generated in task 3.2 is retrieved and the action relative to the non-satisfactory experiments are considered. These actions are the used as input information in the model-based stage to refine the candidate product and generate a new candidate formulation to be verified through the experiment-based stage again. These iterations continue until there is complete agreement between the results of the experiment-based stage and the predictions of the model-based stage.

Sub-task 3.4.1 Comparison between Model-based Calculation and Experimental Results

In this first sub-task, the experimental results generated and collected in sub-task 3.3.2 are compared with the model-based calculations.

In relation to the product property validation experiments, the comparison is purely numerical, as it is desired that the experimental measurements match the predictions performed in sub-task 2.4.1, and it is necessary that they fall into the boundaries of acceptance defined in sub-task 1.3.3. In relation to most of the product performance validation experiments, instead, the output of the experiments can only be "approved" either "not approved", as panel tests are performed in order to validate non-quantifiable product attributes such as color or odor.

By the comparison between the model-based calculations and the experimental results, those experiments that did not validate the predictions are highlighted.

Sub-task 3.4.2 Generation of a List of Refinements

This final sub-task of the integrated methodology is in charge of generating a list of product refinements to be used as an additional input in the model-based stage. These refinements are necessary only in case one or more of the experiments performed in sub-task 3.3.2 do not validate the relative product property and/or performance.

In case non-validations are found in sub-task 3.4.1, the list of actions produced in sub-task 3.2.2 is consulted, in relation to the highlighted experiments, in order to choose a short-list of product refinements. This sub-task is necessary, rather than a simple application of the list of action, because more actions can be recommended for the same reason for non-validation, or two or more reasons for non-validation can have competing suggested actions. Therefore, it is necessary to analyze together all the discrepancies, in order to come out with the optimal list of refinements.

It has to be highlighted here that it is usually recommended to choose refinements involving the latest steps of the model-based stage, rather than the first steps, when possible. This is because a change of an active ingredient, for example, may be followed by the choice of a different solvent mixture (to ensure full solubility, for example), that may then cause different choices of the additives, ending up generating a completely different candidate product, the properties and performances of which need to be verified once more. In case, on the contrary, that no discrepancies appear from the comparison between measurements and predictions, the candidate formulated product is validated.

Table 5.1 summarizes the overall integrated methodology, by highlighting the input, output and necessary tools for each of the sub-tasks describe above, thus giving an exhaustive overview of the data-flow.

Table 5.1 Overview of the data-flow of the integrated methodology, by highlighting input, output and necessary tools for each of the individual methodology sub-task.

Sub-task	Input	Tools	Output
1.1.1	Type of product	Knowledge base	Consumer needs
1.1.2	Consumer needs	Knowledge base	Main and secondary consumer needs
1.2.1	Consumer needs	Knowledge base	Classes of ingredients
1.2.2	Main and secondary consumer needs	Knowledge base	Categories of active ingredients and additives
1.3.1	Consumer needs	Knowledge base	Target properties
1.3.2	Target properties, categories of active ingredients and additives	Knowledge base	Target properties of active ingredients, of additives, of all the ingredients, and of the product
1.3.3	Target properties	Knowledge base	Target values and boundaries of acceptance
2.1.1	Categories of active ingredients, target properties of active ingredients	Structured database, property models	List of candidate active ingredients
2.1.2	List of candidate active ingredients	Structured database, property models	Candidate active ingredients
2.1.3	Candidate active ingredients	Structured database, property models	Candidate active ingredients with properties
2.2.1	Definition of emulsified solvent mixture	Structured database, property models	List of candidate aqueous solvents, organic solvents and surfactant system
2.2.2	Target properties of the product	Knowledge base, property models	Models to be applied in the EMUD algorithm
2.2.3	Target properties of the product	EMUD algorithm	Candidate emulsified solvent mixture
2.2.4	Candidate emulsified solvent mixture	Property models	Candidate emulsified solvent mixture with properties
2.3.1	Categories of additives, target properties of additives, solvent mixture	Structured database, property models	List of candidate additives

2.3.2	List of candidate additives	Structured database, property models	Candidate additives
2.3.3	Candidate additives	Structured database, property models	Candidate additives with properties
2.4.1	Candidate ingredients	Property models	Candidate product with properties
2.4.2	Candidate product	Property models	Stable candidate product
3.1.1	Consumer needs, target properties	Knowledge base	Product properties to be validated
3.1.2	Consumer needs, classes of ingredients	Knowledge base	Product performances to be validated
3.1.3	Product properties and performances to be validated	Knowledge base	List of experiments
3.2.1	List of experiments	Knowledge base	List of possible non validations
3.2.2	Candidate product, list of possible non validations	Knowledge base	List of actions
3.3.1	List of experiments, chemicals and experimental set-up availability	-	Final list of experiments
3.3.2	Final list of experiments	Experimental set-up	Experimental results
3.4.1	Consumer needs, target properties, target values and boundaries of acceptance, experimental results	-	List of non-validations
3.4.2	Candidate product, list of non-validation	Experience, common sense	List of refinements

6

DESIGN
CASE STUDIES

In this chapter, three case studies involving the design of three different consumer products in the emulsified form, are presented.

They consist of:

- A UV-sunscreen (Mattei, Kontogeorgis and Gani, 2012);
- A tank cleaning detergent (Mattei et *al.*, 2014);
- A hand-wash detergent (Mattei et *al.*, 2013).

The methodology illustrated in Figure 5.1 is applied for all the three case studies, but only for the hand-wash detergent (§6.3), the experiment-based stage is fully applied. In fact, the design of the UV sunscreen (§6.1) and of the tank-cleaning detergent (§6.2) involve only tasks 3.1 and 3.2 of stage 3, as the actual experimental work has not been performed.

6.1 UV Sunscreen

The aim of this case study is the design of a UV sunscreen, in the emulsified form, with a high sun protection factor. As mentioned in the introductory paragraph, this conceptual case study consists of stage 1, stage 2 and only part of stage 3 of the methodology presented in chapter 5, as the relative experimental work has been planned, but not performed.

6.1.1 Stage 1: Problem Definition

A UV sunscreen is a product that absorbs or reflects some of the ultraviolet (UV) radiations, and it is applied on the human skin, to help protecting from sunburns, when exposed to sunlight. A large number of sunscreens contain tanning powder to help the skin to

darken or tan; however, tanning powder does not provide protection from UV rays, and the inclusion of such chemicals in the final emulsified product is out of the scope of this case study. However, it has to be noticed that the same methodology highlighted here, can be also applied for the design of such a product.

Sub-task 1.1.1: Collection of the Consumer Needs

From the knowledge-base collected for this case-study (§4.2), it resulted that consumers want a product characterized by the following consumer needs (ψ_i):

- Protection from sunburns;
- Protection from the risk of skin cancer and other similar diseases;
- Prevention from skin-ageing;
- Water-proofness;
- Pleasant odor;
- Pleasant color;
- Pleasant skin feeling;
- Good stability;
- Low toxicity;
- High safety;
- Spray-ability.

Table 6.1 Distinction between main and secondary consumer needs,
for a UV sunscreen in the emulsified form

Consumer Needs (ψ_i)	Main Consumer Needs (ψ_M)	Secondary Consumer Needs (ψ_S)
Protection from sunburns	✓	
Protection from the risk of skin cancer	✓	
Prevention of skin-ageing		✓
Water-proofness		✓
Pleasant odor		✓
Pleasant color		✓
Pleasant skin feeling		✓
Good stability		✓
Low toxicity	✓	
High safety	✓	
Spray-ability		✓

Sub-task 1.1.2: Definition of the Main and the Secondary Consumer Needs

According to the knowledge-base, it is also possible to separate the consumer needs (ψ_i) into main consumer needs (ψ_M) and secondary consumer needs (ψ_S). Table 6.1 illustrates such a distinction.

Sub-task 1.2.1: Collection of the Target Properties I – Not Modelled

In relation to the consumer needs collected in subtask 1.1.1, it is possible to identify eight ingredients (ξ_i), with regards to seven consumer needs, as illustrated in Table 6.2.

Table 6.2 Collection of the classes of ingredients and distinction between active ingredients and additives, for a UV sunscreen in the emulsified form

Consumer Needs (ψ_i)	Classes of Ingredients (ξ_i)	Active Ingredients (ξ_{AI}) or Additives (ξ_{ADD})
Protection from sunburns	UV-A Absorber	Active Ingredient (ξ_{AI})
Protection from the risk of skin cancer	UV-B Absorber	Active Ingredient (ξ_{AI})
	UV Filter	Active Ingredient (ξ_{AI})
Prevention of skin-ageing	Antioxidant	Additive (ξ_{ADD})
Water-proofness	-	-
Pleasant odor	Aroma	Additive (ξ_{ADD})
Pleasant color	Colorant	Additive (ξ_{ADD})
Pleasant skin feeling	Skin conditioner	Additive (ξ_{ADD})
Good stability	Preservative	Additive (ξ_{ADD})
Low toxicity	-	-
High safety	-	-
Spray-ability	-	-

As briefly introduced in §5.2.2, more than one category ingredients can be identified from the analysis of a single consumer need and, vice versa, more than one consumer need can be satisfied by a single class of ingredients. In this specific case study, two consumer needs, the protection from sunburns and the protection from the risk of skin cancer, are merged together as both of them require the protection from the UV radiation. However, two UV absorbers are usually requires, as usually they are effective either against the UV-A radiation (320-400 nm), either against the UV-B radiation (290-320 nm). Moreover, since a sunscreen characterized by a high sun protection factor (SPF) is required, also a UV filter is defined as necessary class of ingredients.

Sub-task 1.2.2: Definition of the Active Ingredients and of the Additives

As from §5.2.2, those ingredients satisfying the main needs ψ_M are classified as active ingredients, while those ingredients satisfying the secondary needs ψ_S are defined as additives. According to the distinction performed in sub-task 1.1.2, illustrated in Table 6.1, the UV-A absorber, the UV-B absorber and the UV filter are the active ingredients of the formulated product, while the antioxidant, the aroma, the colorant and the skin conditioner are the additives, as from Table 6.2.

Sub-task 1.3.1: Collection of the Target Properties II – Modelled

In relation to the consumer needs defined in sub-task 1.1.1, it is possible to collect eight target properties (ζ_i), corresponding to five consumer needs. They are listed in Table 6.3.

Table 6.3 Collection of the target properties and classification of the
target properties, for a UV sunscreen in the emulsified form

Consumer Needs (ψ_i)	Target Properties (ζ_i)	Categories of Target Properties (ζ_i)
Protection from sunburns	-	-
Protection from the risk of skin cancer	-	-
Prevention of skin-ageing	-	-
Water-proofness	Solubility Parameter	ζ_1
Pleasant odor	-	-
Pleasant color	-	-
Pleasant skin feeling	-	-
Good stability	Cloud Point	ζ_4
	Critical Micelle Concentration	ζ_4
	Hydrophilic-Lipophilic Deviation	ζ_4
Low toxicity	Toxicity Parameter	ζ_3
High safety	Flash Point	ζ_3
Spray-ability	Kinematic Viscosity	ζ_4
	Molar Volume	ζ_4

Similarly to what discussed in sub-task 1.2.1, also in relation to the target properties, it is possible that one consumer need requires more than one target properties for its satisfaction, or vice versa. This is the case of the requirement for good stability of the emulsified

product, identifying three target properties: the cloud point, the critical micelle concentration and the hydrophilic-lipophilic balance, and of the requirement for spray-ability, translated into kinematic viscosity and molar volume (or liquid density), as target properties.

Sub-task 1.3.2: Classification of the Target Properties

Out of the eight target properties (ζ_i) collected in sub-task 1.3.1, one of them is classified as relative to the active ingredients (ζ_1), five are relative to the overall product (ζ_4), and two are relative to all the formulation ingredients (ζ_3). None of the target properties collected is classified as relative to the additives (ζ_2). The overall classification is illustrated in Table 6.3.

It has to be underlined that the solubility parameter has been classified as relative only to the active ingredients since it means that only the selected active ingredients must be satisfying the need for water-proofness. It is considered necessary, in fact, that only the fundamental ingredients of the formulation fulfill this requirement.

Sub-task 1.3.3: Identification of the Target Values and of the Boundaries of Acceptance

For each of the target properties retrieve in sub-task 1.3.1, it is necessary to set a target value and/or boundaries of acceptance, so that the relative ingredients can be screened and/or selected on their basis. Table 6.4 lists all the numerical constraints that have been defined for the target properties of the emulsified formulated product.

Table 6.4 Collection of the target values and of the boundaries of acceptance of the target properties, for a UV sunscreen in the emulsified form

Target Properties (ζ_i)	Target Value ($\zeta_{i,T}$)	Boundaries of Acceptance $\zeta_{i,L}$, $\zeta_{i,U}$	Unit of Measure
Solubility Parameter	-	$\delta_T < 25$	$MPa^{1/2}$
Cloud Point	-	$CP > 70$	°C
Critical Micelle Concentration	-	$c_{surf} > CMC_{surf}$	mol/m^3
Hydrophilic-Lipophilic Deviation	-	$HLD \neq 0$	[-]
Toxicity Parameter	-	$LC_{50} > 3.16$	mol/m^3
Flash Point	-	$T_f > 70$	°C
Kinematic Viscosity	-	$v < 75$	cS
Molar Volume	-	$100 < V_m < 150$	l/kmol

6.1.2 Stage 2: Model-based Stage

In this stage, computer-aided tools are used in order to screen between several candidate ingredients, in order to propose a candidate emulsified formulation, to be further validated or refined in the experiment-bases stage. The information collected in stage 1 (§6.1.1) are used as input to the work-flow reported in Figure 5.3.

Sub-task 2.1.1: Screening of the Candidate Active Ingredients

Three databases are retrieved in this sub-task: the UV-A absorbers database, the UV-B absorbers database, and the UV filter database. According to the European regulation on cosmetic products, however, only a limited number of chemicals can be included in sunscreen formulations, as active ingredients, and constraints on the maximum concentration allowed are also set (EC Regulation n.1223-2009). All the 27 chemicals permitted are available in our databases, and those not mentioned by the European regulation are discarded.

Moreover, the constraints on the Hildebrand solubility parameter, on the toxicity parameter and on the flash point reported in Table 6.4 are applied. Short-lists of UV-A and UV-B absorbers, as well as of UV filters, are generated.

Sub-task 2.1.2: Selection of the Best Active Ingredients

In order to properly choose the best between the candidate active ingredients generated in sub-task 2.1.1, it is first necessary to define a performance index, on the basis of which the selection can be performed.

As mentioned in §5.3.1, the performance index usually consists of one or more of the following criteria:
- Effectiveness;
- Safety and health;
- Environment;
- Cost.

In this specific case, as a UV sunscreen characterized by a high sun protection factor is the aim of the case study, the effectiveness is chosen as performance index. The effectiveness of the candidate active ingredients, alone and in mixtures (Couteau et al., 2009) has been considered and on this basis, Avobenzone, Octyl Salicylate and Zinc Oxide have been selected as the best UV-A absorber, UV-B absorber and UV filter, respectively. Figure 6.1 illustrates the chemical structure of the chosen active ingredients.

Figure 6.1 Chemical structures of avobenzone (left) and octyl salicylate (right), chosen active ingredients for the UV sunscreen in the emulsified form

Sub-task 2.1.3: Collection of the Active Ingredients Properties

For each of the active ingredients selected in sub-task 2.1.2, it is necessary to retrieve from the relative databases (§4.1.1) all the thermo-physical properties that are required in the further steps of the methodology. In case these are not available, pure component property models are used to predict them. In relation to the three active ingredients chosen in this case study, the results are collected in Table 6.5.

Table 6.5 Candidate Active Ingredients with thermo-physical properties, for a UV sunscreen in the emulsified form

Commercial Name	Hildebrand Sol. Par. $[MPa^{1/2}]$	MW $[g/mol]$	Density $[g/ml]$	Liquid Viscosity $[cP]$	Flash Point $[°C]$	Toxicity Parameter $[mol/m^3]$	Cost $[\$/kg]$
Avobenzone	19.4	310.4	1.04	65.0	>100	5.83	321.0
Octyl Salicylate	20.4	250.3	1.01	69.0	>100	5.54	80.9
Zinc Oxide	-	81.4	5.61	-	>100	3.93	63.4

Sub-task 2.2.1: Retrieval of the Databases

UV sunscreen in the emulsified form are usually oil-in-water emulsions; therefore, three different databases need to be retrieved: one in relation to the aqueous solvent phase, one in relation to the organic solvent phase, and one in relation to the surfactant system.

The aqueous solvent phase usually consists of water. This is due to the fact that a UV sunscreen needs to be directly applied on the skin, therefore other aqueous solvents (such as, water soluble alcohols) are not considered. Moreover, from an economic point of view, water is extremely convenient.

The organic solvent phase, instead, usually consists of an ester. These chemicals, in fact, are very widely used for personal care, as well as pharmaceutical products, as they have an extremely wide solubility range, they are recognized being non-toxic and they are not miscible with water, thus able to generate an emulsion if a proper surfactant system is chosen.

Finally, the non-ionic surfactants database is retrieved, in relation to the surfactant system, as non-ionic surfactants are usually preferred for those consumer products where the production of foam in unwanted, and as they are characterized by enhanced stability, in particular against electrolytes.

Sub-task 2.2.2: Choice of the Property Models

Appropriate models for the thermo-physical target properties to be calculated in the EMUD algorithm need to be chosen in this sub-task. However, as it has been already mentioned in §5.3.2, mixture models characterized by the linear mixing rule are preferred at this level of the methodology, for calculating the properties of each of the two phases. On the other hand, for the calculation of the overall properties of the emulsified solvent mixture, dedicated models are necessary (§3.4), and therefore the model selection is the same described in Figure 3.4.

Sub-task 2.2.3: Application of the EMUD Algorithm

The algorithm for the design of emulsified solvent mixtures is applied here, with the database and property models chosen in sub-tasks 2.2.1 and 2.2.2. The property constraints set in the EMUD algorithm are listed below, divided in relation to the calculation levels:

- Pure component properties

  Aqueous solvent phase: $LC_{50} > 3.16$ mol/m^3
  $T_f > 70°C$

  Organic solvent phase: $LC_{50} > 3.16$ mol/m^3
  $T_f > 70°C$
  $17.4 < \delta_T < 21.4$ MPa$^{1/2}$

  Surfactant mixture: $LC_{50} > 3.16$ mol/m^3
  $T_f > 70°C$
  $T_C > 70°C$

- Mixture properties with linear mixing rule

  $$60 < V_m < 150 \text{ l/kmol}$$
  $$C_{surf} \geq 10 \cdot CMC_{surf}$$

- Mixture properties with non-linear mixing rule

  $$\nu < 75 \text{ cS}$$

- Stability check

  $$HLD \neq 0$$

Given the above constraints, the candidate emulsified solvent mixture generated from the EMUD algorithm consists of:

- Water (aqueous solvent phase): 62.8 % in weight;
- Butyl acetate (organic solvent phase): 37.1% in weight;
- Octyl esaethylene oxide (surfactant system): 0.1% in weight.

Their thermo-physical pure properties are given in Table 6.6.

Sub-task 2.2.4: Collection of the Emulsified Solvent Mixture Properties

Rigorous models are applied to the candidate emulsified solvent mixture determined in sub-task 2.2.3 and the results are compared with the boundaries on the target properties set in sub-task 1.3.3. The results are summarized below:

- Kinematic viscosity: 0.95 cS (lower boundary: 0; upper boundary: 75 cS);
- Molar volume: 60.7 l/kmol (lower boundary: 60; upper boundary: 150 l/kmol);
- Hydrophilic-lipophilic deviation: -2.5 (lower boundary: 0).

The results obtained with the rigorous models are in good agreement with the boundaries that have been previously set.

Table 6.6 Candidate solvent mixture ingredients with thermo-physical properties

Commercial Name	Solubility Parameter [MPa$^{1/2}$]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [˚C]	Toxicity Parameter [mol/m^3]	Cost [$/kg]
Butyl Acetate	18.4	116.2	0.88	0.86	>100	3.60	74
Water	30.1	18.0	1.00	0.89	-	-	-

Comm. Name	CMC [mol/L]	CP [˚C]	σ [mN/m]	HLB [-]	MW [g/mol]	ρ [g/ml]	μ [cP]	T$_f$ [˚C]	LC$_{50}$ [mol/m^3]	Cost [$/kg]
Octyl Esaethylene Oxide	0.009	73	29	13.4	395	1.07	36	>100	3.97	175

Sub-task 2.3.1: Screening of the Candidate Additives

As from the information collected in Table 6.2, five databases are retrieved in this sub-task: the antioxidant database, the aroma database, the colorant database, the preservative database and the skin conditioner database.

Constraints on the Hildebrand solubility parameter, on the toxicity parameter and on the flash point reported in are applied:

- Solubility parameter $\quad\quad 15.4 < \delta_T < 21.4$ MPa$^{1/2}$ (organic phase)

  $27.1 < \delta_T < 33.1$ MPa$^{1/2}$ (aqueous phase)

- Flash point $\quad\quad\quad\quad\quad$ T$_f$ > 70°C
- Toxicity parameter $\quad\quad$ LC$_{50}$ > 3.16 mol/m^3

Short-lists of candidate additives are generated.

Sub-task 2.3.2: Selection of the Best Additives

Similarly to sub-task 2.1.2, the selection of the best additives is carried out on the basis of a performance index consisting of one or more of the following criteria: effectiveness, safety, health, environment and cost. In this specific case study, the minimization of the

cost of the raw materials is identified as a problem objective, therefore the cheapest candidates of the short-lists generates in sub-task 2.3.1 are selected.

α-tocopherol is chosen among the antioxidants, laevo-menthol as an aroma and a skin conditioner, and heptylparaben as a preservative. No additive has been chosen as a colorant, because of the selection of zinc oxide as one of the active ingredients. Zinc oxide, in fact, is dispersed in the formulation, rather than dissolved, and therefore the color of the product needs to be white.

Figure 6.2 illustrates the chemical structure of the chosen additives.
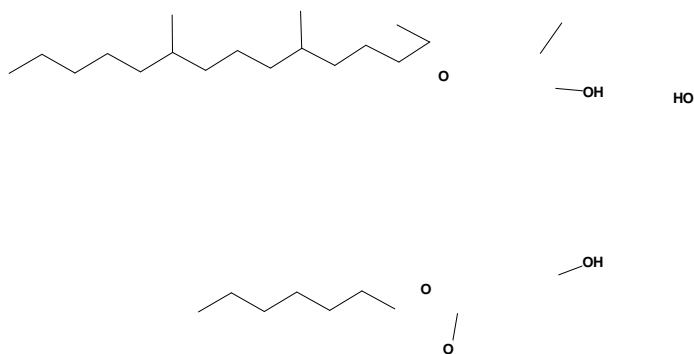
Figure 6.2 Chemical structures of α-tocopherol (top-left), laevo-menthol (top right), and heptylparaben (bottom), chosen additives of the UV sunscreen in the emulsified form

Sub-task 2.3.3: Collection of the Additives Properties

The thermo-physical properties of all the candidate additives are retrieved from the databases, or estimated with pure component property models. The results are collected in Table 6.7.

Table 6.7 Candidate Additives with thermo-physical properties, for a UV sunscreen in the emulsified form

Commercial Name	Hildebrand Sol. Par. [MPa$^{1/2}$]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [˚C]	Toxicity Parameter [mol/m^3]	Cost [$/kg]
α-tocopherol	15.6	430.7	0.95	224.2	>100	8.62	927.0
laevo-menthol	15.9	156.3	0.98	12.1	>100	3.92	362.8
Heptylparaben	21.2	236.3	1.02	101.9	>100	5.40	112.9

Sub-task 2.4.1: Calculation of the Product Properties

After the addition of the additives, rigorous mixture property models are applied to the candidate overall formulation, and the results are compared once more with the boundaries on the target properties set in sub-task 1.3.3. The results are summarized below:

- Kinematic viscosity: 1.11 cS (lower boundary: 0; upper boundary: 75 cS);
- Molar volume: 69.2 l/kmol (lower boundary: 60; upper boundary: 150 l/kmol);

The results obtained with the rigorous models are in agreement with the boundaries that have been previously set.

Sub-task 2.4.2: Evaluation of the Product Stability

As a last sub-task of the model-based stage, the stability of the product as an emulsion has to be checked with the HLD approach. The same correlation was employed in the EMUD algorithm, and the results (HLD = -2.5) guarantee the stability of the emulsified solvent mixture. Here, the calculation has to be repeated, considering that some ingredients between active ingredients and additives can influence the HLD-value. In this particular case, octyl salicylate, α-tocopherol, laevo-menthol and heptylparaben are alcohols and they therefore contribute in changing the HLD-value by influencing the parameter A of equation 3.34. However, the overall effect is advantageous for the product stability, as the new HLD-value calculate in this sub-task is -3.1.

The candidate formulation is the validated by the model-based stage and it is accepted for experiment-based validation in the third stage of the integrated methodology.

The candidate emulsified formulation to be verified and/or refined in the experiment-based stage is given in Table 6.8.

Table 6.8 Candidate formulation for a UV sunscreen in the emulsified form

	Class of Ingredient	Commercial Name	Weight Percentage
Active Ingredients	UV-A absorber	Avobenzone	0.99%
	UV-B absorber	Octyl salicylate	1.22%
	UV filter	Zinc oxide	8.56%
Solvent Mixture	Aqueous solvent	Water	54.9%
	Organic solvent	Butyl acetate	32.4%
	Surfactant	Octyl esaethylene oxide	0.09%
Additives	Antioxidant	α-tocopherol	0.34%
	Aroma and skin conditioner	laevo-menthol	0.59%
	Preservative	Heptylparaben	0.91%

6.1.3 Stage 3: Experiment-based Stage

In this stage, the knowledge base is used to generate a list of experiments to be made in order to validate the product properties and performances, and then a list of actions to be performed in case any of the experiments does not validate. In relation to this case study, tasks 3.3 and 3.4 have not been considered.

Sub-task 3.1.1: Collection of the Product Properties to be validated

In this sub-task, the product properties to be validated with experiments are collected. They consist of the target properties II collected in sub-task 1.3.1 and they are reported below, together with their classification:
- Solubility parameters (ζ_1);
- Cloud point (ζ_4);
- Critical micelle concentration (ζ_4);
- Hydrophilic-lipophilic deviation (ζ_4);
- Toxicity parameter (ζ_3);
- Flash point (ζ_3);
- Kinematic viscosity (ζ_4);
- Molar volume (ζ_4).

Sub-task 3.1.2: Collection of the Product Performances to be validated

In this sub-task, the product performances to be validated with experiments are collected, and they consist of the target properties I that have been defined in sub-task 1.2.1. They are listed below:
- Protection from sunburns;
- Protection from the risk of skin cancer;
- Prevention of skin-ageing;
- Water-proofness;
- Pleasant odor;
- Pleasant color;
- Pleasant skin feeling;
- Good stability;
- Low toxicity;
- High safety;
- Spray-ability.

Sub-task 3.1.3: Generation of a List of Experiments

In this sub-task, a list of experiments for the validation of the product properties collected in sub-task 3.1.1 and of the product performances collected in 3.1.2 is generated. Table 6.9 gives the list of proposed experiments, together with the relative product properties and product performances to be validated.

Table 6.9 List of experiments for a UV sunscreen in the emulsified form

Consumer Need	Product Property Validation Experiment	Product Performance Validation Experiment
Protection from sunburns	-	Measurement of the sun protection factor
Protection from the risk of skin cancer	-	
Prevention of the skin-ageing	-	-
Water-proofness	-	-
Pleasant odor	-	Panel test for the odor
Pleasant color	-	Panel test for the color
Pleasant skin feeling	-	Measurement of pH Panel test for the skin feeling
Good stability	Measurement of the phase inversion temperature	Solubility test Solvent mixture stability test Shelf-life test
Low toxicity	-	-
High safety	-	-
Spray-ability	Measurement of the viscosity Measurement of the molar volume	Panel test for the spray-ability

Sub-task 3.2.1: Collection of the possible Non-validations

In this sub-task, all the possible reasons for the experiments listed in the previous sub-task to be non successful are collected. Heuristics and correlations linking the product properties and performances with variables such as the type and the concentration of certain ingredients are used. The information collected in this sub-task are fundamental for a reliable generation of a list of actions. As a table containing all the possible non-validations is considered redundant with the table containing the list of actions, it is preferred to show only the second.

Sub-task 3.2.2: Generation of a List of Actions

In this sub-task, a list of actions corresponding to each of the experiments for the validation of the product properties and performances is generated. Actions influencing the latest steps of the model-based methodology are preferred so that, for example, changing an

additive is usually recommended over modifying the solvent mixture. Table 6.10 gives the list of actions, in relation to the validation experiments proposed in sub-task 3.1.3.

Table 6.10 List of actions for a UV sunscreen in the emulsified form

Validation Experiment	Action 1	Action 2
Measurement of the PIT	Add an adequate additive	Change the solvent mixture
Measurement of v	Change the droplet size distribution	Add an adequate additive
Measurement of V_m	Add an adequate additive	Change the solvent mixture
Measurement of the SPF	Change the concentration of the active ingredients	Change the solvent mixture
Panel test for the odor	Add an adequate additive	Change the additive
Panel test for the color	Add an adequate additive	Change the additive
Measurement of pH	Add an adequate additive	Change the additive
Panel test for the skin feeling	Add an adequate additive	Change the solvent mixture
Solubility test	Change the solvent mixture	-
Solvent mixture stability test	Change the solvent mixture	-
Shelf-life test	Change the solvent mixture	-
Panel test for the spray-ability	Change the solvent mixture	-

6.2 Tank Cleaning Detergent

The aim of this second case study is the design of a tank cleaning detergent, in the emulsified form, for the removal of palm oil. This is also a conceptual case study, meaning that only a part of the experiment-based stage is presented, as the experimental work has been planned, but not performed.

6.2.1 Stage 1: Problem Definition

Commercial and industrial detergents have recently become extremely sophisticated, in order to address a broad range of cleaning tasks and to deliver superior performances with a minimum of effort an time. These products, by definition, consist of different chemicals, each with a specific function related to the needs of the product: surfactants, builders, bleaching agents, enzymes and minors, usually mixed together with a carrier, necessary to keep the blend in the desired physical form.

Sub-task 1.1.1: Collection of the Consumer Needs

From the knowledge-base collected for this case-study (§4.1), it resulted that consumers want a product characterized by the following consumer needs (ψ_i):

- Wetting of the substrate;
- Dissolution of the dirt;
- Suspension of the dirt;
- Low foam-ability;
- Good stability to temperature and dilution;
- Good stability to electrolytes;
- Low impact on the substrate;
- High dissolution of the dirt;
- Visible color;
- Easy application;
- Low toxicity;
- High safety.

Among these consumer needs, the first three are usually merged together under the denomination of "cleaning performances". However, as each of the three actions requires different target properties, they have been listed singularly.

Sub-task 1.1.2: Definition of the Main and the Secondary Consumer Needs

According to the knowledge-base, it is also possible to distinguish the main consumer needs (ψ_m) from the secondary consumer needs (ψ_s), as illustrated in Table 6.11.

Table 6.11 Distinction between main and secondary consumer needs
for a tank-cleaning detergent in the emulsified form

Consumer Needs (ψ_i)	Main Consumer Needs (ψ_M)	Secondary Consumer Needs (ψ_S)
Wetting of the substrate	✓	
Dissolution of the dirt	✓	
Suspension of the dirt	✓	
Low foam-ability	✓	
Good stability to temperature and dilution	✓	
Good stability to electrolytes		✓
Low impact on the substrate		✓
High dissolution of dirt		✓
Visible color		✓
Easy application		✓
Low toxicity	✓	
High safety	✓	

Sub-task 1.2.1: Collection of the Target Properties I – Not Modelled

Six classes of ingredients (ξ_i) are necessary in the emulsified formulated product, in order to satisfy the target properties that are not modelled for this case study. Table 6.12 provides a short list of them.

Table 6.12 Collection of the classes of ingredients and distinction between active ingredients and additives, for a tank-cleaning detergent in the emulsified form

Consumer Needs (ψ_i)	Class of Ingredients (ξ_i)	Active Ingredients (ξ_{AI}) or Additives (ξ_{ADD})
Wetting of the substrate		
Dissolution of the dirt	Non-Ionic Surfactant	Active Ingredient (ξ_{AI})
Suspension of the dirt		
Low foam-ability		
Good stability to temperature and dilution	Co-surfactant	Additive (ξ_{ADD})
Good stability to electrolytes	Softener	Additive (ξ_{ADD})
Low impact on the substrate	Builder	Additive (ξ_{ADD})
High dissolution of dirt	Bleaching Agent	Additive (ξ_{ADD})
Visible color	Colorant	Additive (ξ_{ADD})
Easy application	-	-
Low toxicity	-	-
High safety	-	-

Similarly to what has been underlined in the previous case study, also here the presence of one class of ingredients (the non-ionic surfactants) is considered necessary to satisfy more than one consumer need. In this specific case, all the needs related to the cleaning performances can be satisfied with the presence of surfactants, while the specific need for low foam-ability limits the search-space to the non-ionic surfactants.

Sub-task 1.2.2: Definition of the Active Ingredients and of the Additives

The distinction between active ingredients and additives strictly follows the classification of the main needs ψ_m and secondary needs ψ_s. According to such a classification, then, non-ionic surfactants are identified as the active ingredients of this product, while a softener, a builder, a bleaching agent and a colorant are necessary in the emulsified formulation as additives. Table 6.12 illustrates the distinction above.

Sub-task 1.3.1: Collection of the Target Properties II – Modelled

In relation to the consumer needs defined in sub-task 1.1.1 and reported in Table 6.11, also a set of target properties that are modelled (that is, target properties II) are necessary. In this specific case study, nine target properties (ζ_i) are collected and they are reported in Table 6.13.

Sub-task 1.3.2: Classification of the Target Properties

Out of the eight target properties (ζ_i) collected in sub-task 1.3.1, five of them are classified as relative to the active ingredients (ζ_1), two are relative to the overall product (ζ_4), and two are relative to all the formulation ingredients (ζ_3). None of the target properties collected is classified as relative to the additives (ζ_2). The overall classification is illustrated in Table 6.13.

Table 6.13 Collection of the target properties and classification of the target properties, for a tank-cleaning detergent in the emulsified form

Consumer Needs (ψ_i)	Target Properties (ζ_i)	Categories of Target Properties (ζ_i)
Wetting of the substrate	Surface Tension	ζ_1
Dissolution of the dirt	Solubility Parameters	ζ_1
Suspension of the dirt	Hydrophilic-Lipophilic Balance	ζ_1
Low foam-ability	Surface Tension	ζ_1
	Critical Micelle Concentration	ζ_1
Good stability to temperature and dilution	Cloud Point	ζ_1
	Hydrophilic-Lipophilic Deviation	ζ_4
Good stability to electrolytes	-	-
Low impact on the substrate	-	-
High dissolution of dirt	-	-
Visible color	-	-
Easy application	Viscosity	ζ_4
Low toxicity	Toxicity Parameter	ζ_3
High safety	Flash Point	ζ_3

It is necessary to highlight the fact that the surface tension is strictly a property of the mixture, and therefore it should be classified as ζ_4. However, as reported in §3.4.2, the

surface tension of an oil-in-water emulsion can be considered a pure property of the sur-factant. Therefore, as non-ionic surfactants have been identified as the active ingredients, this target property is classified as relative to the active ingredient (ζ_1).

Sub-task 1.3.3: Identification of the Target Values and of the Boundaries of Acceptance

For each of the target properties retrieve in sub-task 1.3.1, a target value and/or bounda-ries of acceptance have to be set, for ingredient screening and selection reasons. Table 6.14 lists all the numerical constraints defined for the collected target properties.

Table 6.14 Collection of the target values and of the boundaries of acceptance of the target properties, for a tank-cleaning detergent in the emulsified form

Target Properties (ζ_i)	Target Value ($\zeta_{i,T}$)	Boundaries of Acceptance $\zeta_{i,L}$, $\zeta_{i,U}$	Unit of Measure
Surface Tension	-	$\sigma < 40$	mN/m
Solubility Parameters	$\delta_D = 17.7$ $\delta_P = 3.5$ $\delta_H = 3.7$	$15.7 < \delta_D < 19.7$ $1.5 < \delta_P < 5.5$ $1.7 < \delta_H < 5.7$	$MPa^{1/2}$ $MPa^{1/2}$ $MPa^{1/2}$
Hydrophilic-Lipophilic Balance	-	$HLB > 12$	[-]
Surface Tension	-	$\sigma > 25$	mN/m
Critical Micelle Concentration	-	$CMC > 0.01$	mol/l
Cloud Point	-	$CP > 80$	°C
Hydrophilic-Lipophilic Deviation	-	$HLD \neq 0$	[-]
Viscosity	-	$v < 25$	cS
Toxicity Parameter		$LC_{50} > 3.16$	mol/m^3
Flash Point		$T_f > 70$	°C

It needs to be noticed that the target values relative to the solubility parameters consist of the Hansen solubility parameter of the palm oil. Moreover, two different boundaries of acceptance are set for the surface tension, in order to underline that different consumer needs may require different constraints on the same target property, as in this specific case.

6.2.2 Stage 2: Model-based Stage

In this stage, a candidate emulsified formulation is generated, by means of models, struc-tured databases, and dedicated algorithm. The input of this stage consists of the infor-mation collected in the previous stage, in terms of classes of ingredients and target prop-erties.

Sub-task 2.1.1: Screening of the Candidate Active Ingredients

It is necessary to retrieve only one database in this sub-task, for this specific case study: the non-ionic surfactants database. Constraints on the surface tension, on the Hansen solubility parameters, on the hydrophilic-lipophilic balance, on the critical micelle concentration, on the cloud point, on the toxicity parameters and on the flash point, as give in Table 6.14 are applied and a short-list of candidate non-ionic surfactants s generated.

Sub-task 2.1.2: Selection of the Best Active Ingredients

Similarly to what described in relation to the previous case study, the effectiveness is chosen as the criteria for the selection of the best active ingredient, over safety, health, environment and cost. In this specific case, the effectiveness of the active ingredient is identified with the target values set for the Hansen solubility parameters in sub-task 1.3.3. In case two or more candidate non-ionic surfactants are characterized by comparable effectiveness, the cheapest solution is chosen.

Decyl esaethylene oxide is chosen as the best active ingredient for this case study; its chemical structure is given in Figure 6.3.
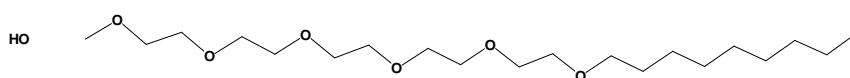
Figure 6.3 Chemical structure of decyl esaethylene oxide, chosen as active ingredient for a tank-cleaning detergent in the emulsified form

Sub-task 2.1.3: Collection of the Active Ingredients Properties

In this sub-task, it is necessary to retrieve from the non-ionic surfactants database all the thermo-physical properties that are required in the further steps of the methodology. Pure component property models are used to predict those that are not available. In relation to the active ingredients chosen in this case study, the results are collected in Table 6.15.

Table 6.15 Candidate active ingredient with thermo-physical properties, for a tank-cleaning detergent in the emulsified form

Comm. Name	CMC [mol/L]	CP [°C]	σ [mN/m]	HLB [-]	MW [g/mol]	ρ [g/ml]	μ [cP]	T_f [°C]	LC_{50} [mol/m³]	Cost [$/kg]
Decyl Esaethylene Oxide	0.013	80.3	33	12.6	422.59	0.99	47	>100	4.55	127

Sub-task 2.2.1: Retrieval of the Databases

Tank-cleaning detergents are usually oil-in-water emulsions, therefore three databases are required: one for the aqueous solvent phase, one for the organic solvent phase, and one for the surfactant system. However, in this specific case study, a surfactant has already

been chosen as the active ingredient of the emulsified formulation; therefore, there is no need for selecting another surfactant as part of the emulsified solvent mixture. In relation to the organic solvent phase, moreover, from the problem definition palm oil is defined as the chemical to be removed from the substrate during the cleaning process and therefore there is no need for retrieving any database relative to organic solvent phase chemicals, as it consists of palm oil.

Finally, the aqueous solvent phase is usually water as it is extremely cheap and it can be freely discharged after use.

Sub-task 2.2.2: Choice of the Property Models

Also in this case-study, in the EMUD algorithm, the mixture property model characterized by linear mixing rules is preferred, in relation to the molar volume, while the non-linear mixture property model for the viscosity has to be used.

Sub-task 2.2.3: Application of the EMUD Algorithm

In this specific case, as the chemicals constituting each of the three phases are defined a priori, the first level of the EMUD algorithm is not necessary. The property constraints set in the other level of the algorithm are listed below:

- Mixture properties with linear mixing rule
$$C_{surf} \geq 10 \cdot CMC_{surf}$$
- Mixture properties with non-linear mixing rule
$$\nu < 25 \text{ cS}$$
- Stability check
$$HLD \neq 0$$

Under the constraints given above, the composition of the emulsified solvent mixture is generated by minimizing its cost. Moreover, the molar ratio between water and palm oil is set to a constant value, determined by processing constraints. This value is usually set to 7 (Anonymous, 2012).

The candidate solvent mixture consists, then, of:

- Water (aqueous solvent phase): 84.6 % in mole;
- Decyl esaethylene oxide (surfactant system): 15.4% in mole.

It is necessary to underline that the property constraints above are applied to the formulated product during the application, that is, in presence of palm oil. In these conditions, the composition of the emulsified solvent mixture is as follows:

- Water: 75.5 % in mole;
- Palm oil: 10.8% in mole;
- Decyl esaethylene oxide: 13.7% in mole.

The thermo-physical properties considered in this case study for the palm oil are reported in Table 6.16.

Table 6.16 Thermo-physical properties of palm oil

Commercial Name	Solubility Parameter [MPa$^{1/2}$]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [˚C]	Toxicity Parameter [mol/m^3]	Cost [$/kg]
Palm Oil	18.4	269.4	0.93	77.2	>100	-	-

Sub-task 2.2.4: Collection of the Emulsified Solvent Mixture Properties

In this sub-task, rigorous models are applied to the candidate solvent mixture to verify that the target properties match the constraints set in the problem definition stage. The results are given below:

- Kinematic viscosity: 5.30 cS (upper boundary: 25 cS);
- Hydrophilic-lipophilic deviation: -5.1 (lower boundary: 0).

Agreement is found between the properties calculated with rigorous models, and the boundaries previously set.

Sub-task 2.3.1: Screening of the Candidate Additives

According to Table 6.12, five additives are necessary to satisfy the secondary need of this case studies, and therefore five databases need to be retrieved: the bleaching agents database, the builders database, the colorants database, the co-surfactants database and the softeners database.

Constraints are set in terms of solubility parameters (or qualitative solubility, when the solubility parameters are not available), toxicity parameter and flash point, as below:

- Solubility parameter $\quad\quad\quad$ $27.1 < \delta_T < 33.1$ MPa$^{1/2}$

  soluble in water
- Flash point $\quad\quad\quad\quad\quad\quad$ $T_f > 70°C$
- Toxicity parameter $\quad\quad\quad$ $LC_{50} > 3.16$ mol/m^3

Short-lists of candidate additives are generated accordingly.

Sub-task 2.3.2: Selection of the Best Additives

Also in this case study, the selection criteria for the additives is the cost, that is, the cheapest chemical per each class of ingredients that match the constraints given in the previous sub-task is selected. As a consequence, the selected additives are:

- Bleaching agent: Potassium hydroxide
- Builder: Sodium silicate
- Colorant: Titanium oxide
- Co-surfactant: Butyl glycol
- Softener: Sodium triphosphate

Figure 6.4 illustrates the chemical structure of the chosen additives.
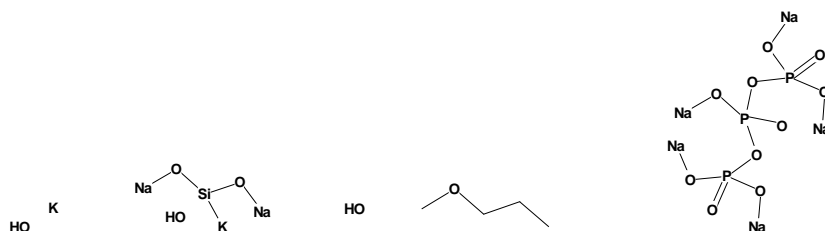
Figure 6.4 Chemical structures of, from the left, potassium hydroxide, sodium silicate, butyl glycol and sodium triphosphate, chose as additives for a tank-cleaning detergent in the emulsified form

Table 6.17 Candidate additives with thermo-physical properties, for a tank-cleaning detergent in the emulsified form

Commercial Name	Qualitative Solubility [-]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [°C]	Toxicity Parameter [mol/m³]	Cost [$/kg]
Potassium Hydroxide	-	56.11	2.12	-	-	3.19	57
Sodium Silicate	aqueous	122.06	2.61	-	-	3.46	68
Titanium Oxide	-	63.87	4.95	-	-	3.32	71
Butyl Glycol	aqueous	118.17	0.90	3.67	75	3.64	24
Sodium Triphosphate	aqueous	367.86	2.52	-	-	4.06	134

Sub-task 2.3.3: Collection of the Additives Properties

For each of the selected additives, the relevant thermo-physical properties are retrieved from the databases, or estimated with property models, if necessary. The results are collected in Table 6.17.

Sub-task 2.4.1: Calculation of the Product Properties

Rigorous mixture property models are applied once more to verify that the presence of the active ingredients and of the additives does not cause the product target properties not to match the constraints defined in the problem definition stage.

The result of the calculation is summarized below:

- Kinematic viscosity: 7.21 cS (upper boundary: 25 cS).

As it has been highlighted in the previous sub-task, information about the contribution to the overall viscosity of part of the additives are missing. However, the calculated value of the kinematic viscosity is enough lower than the upper boundary to give confidence

that the inclusion of the additives would not cause the product property to be out of the constraint.

Sub-task 2.4.2: Evaluation of the Product Stability

In this last sub-task of the model-based stage, the HLD approach is applied to the overall product to check its stability as an emulsion. The presence of several dispersed as well as dissolved species is expected to alter the HLD-value calculated in relation to the solvent mixture.

Table 6.18 Candidate formulation for a tank-cleaning detergent in the emulsified form

	Class of Ingredient	Commercial Name	Mole Percentage
Active Ingredients	Non-ionic Surfactant	Decyl esaethylene oxide	12.3
Solvent Mixture	Aqueous solvent	Water	67.9
Additives	Bleaching agent	Potassium Hydroxide	8.4
	Builder	Sodium Silicate	1.9
	Colorant	Titanium Oxide	3.4
	Co-surfactant	Butyl Glycol	3.6
	Softener	Sodium Triphosphate	2.5

The new HLD-value calculated in this sub-task is -2.6, showing an increase compared to the emulsified solvent mixture, but still low enough to ensure stability.

The candidate formulation has then been validated in the model-based stage and it is given in details in Table 6.18.

6.2.3 Stage 3: Experiment-based Stage

The aim of this third stage is to validate and/or refine by means of experiments the candidate formulation given in Table 6.18. In relation to this case study, only task 3.1 and 3.2 are performed, that is, no experimental work has taken place.

Sub-task 3.1.1: Collection of the Product Properties to be validated

The product properties to be validated with experiments are collected in this sub-task. They represent those target properties defined as target properties II that have been collected in sub-task 1.3.1. They are listed below, according to their classification:

- Surface tension (ζ_1);

- Solubility parameters (ζ_1);
- Hydrophilic-lipophilic balance (ζ_1);
- Critical micelle concentration (ζ_1);
- Cloud point (ζ_1);
- Hydrophilic-lipophilic deviation (ζ_4);
- Kinematic viscosity (ζ_4);
- Toxicity parameter (ζ_3);
- Flash point (ζ_3).

Sub-task 3.1.2: Collection of the Product Performances to be validated

The product performances to be validated with experiments are collected in this sub-task. They represent those target properties defined as target properties I that have been defined in sub-task 1.2.1. They are listed below:

- Wetting of the substrate;
- Dissolution of the dirt;
- Suspension of the dirt;
- Low foam-ability;
- Good stability to temperature and dilution;
- Good stability to electrolytes;
- Low impact on the substrate;
- High dissolution of dirt;
- Visible color;
- Easy application;
- Low toxicity;
- High safety.

Sub-task 3.1.3: Generation of a List of Experiments

On the basis of the product properties and performances to be validated that have been collected in the previous sub-tasks, here a proper list of relative experiments to be performed is generated. Table 6.19 lists them, next to the product property or performance.

Sub-task 3.2.1: Collection of the possible Non-validations

As in the previous case study, the reasons for which the measured product properties and performances may not match the predictions are collected in this sub-task but they are not shown as these information are better explained by means of the list of actions generated in the next sub-task.

Table 6.19 List of experiments for a tank-cleaning detergent in the emulsified form

Consumer Need	Product Property Validation Experiment	Product Performance Validation Experiment
Wetting of the substrate	Measurement of the surface tension	Panel test for the wetting
Dissolution of the dirt	-	Panel test for the dissolution
Suspension of the dirt	-	Panel test for the suspension
Low foam-ability	Measurement of the surface tension	Panel test for the foam-ability
Good stability to temperature and dilution	Measurement of the phase inversion temperature	Solubility test Solvent mixture stability test Shelf-life test
Good stability to electrolytes	-	Solvent mixture stability test
Low impact on the substrate	-	-
High dissolution of dirt	-	-
Visible color	-	Panel test for the color
Easy application	Measurement of the viscosity	-
Low toxicity	-	-
High safety	-	-

Sub-task 3.2.2: Generation of a List of Actions

In this sub-task, each of the experiments listed in sub-task 3.1.3 is coupled with a corresponding action, or amendment, to be taken in case there is no match between the measurement and the prediction. Table 6.20 provides the list of actions relative to the list of experiments generated for this case study, as in table 6.19.

Table 6.20 List of actions for a tank-cleaning detergent in the emulsified form

Validation Experiment	Action 1	Action 2
Measurement of σ	Change the solvent mixture	Change the active ingredient
Measurement of the PIT	Add an adequate additive	Change the solvent mixture
Measurement of v	Change the droplet size distribution	Add an adequate additive
Panel test for wetting	Change the solvent mixture	Change the active ingredient
Panel test for the dissolution	Change the concentration of the active ingredient	Change the active ingredient
Panel test for the suspension	Change the concentration of the active ingredient	Change the solvent mixture
Panel test for the foam-ability	Change the concentration of the active ingredient	Change the active ingredient
Solubility test	Change the solvent mixture	-
Solvent mixture stability test	Change the solvent mixture	-
Shelf-life test	Change the solvent mixture	-
Panel test for the color	Add an adequate additive	Change the additive

6.3 Hand-Wash Detergent

This third and last case study focuses in the design of a hand-wash, in the emulsified form, with emollient action. This is the first and only complete case study developed in this work, meaning that the whole methodology has been applied and that the experimental work has been planned and performed accordingly.

6.3.1 Stage 1: Problem Definition

Commercial hand-wash, as well as industrial detergents, are nowadays tailor-made designed in order to satisfy a wide variety of consumer assessments. That is, the ingredients of the formulation are carefully selected in order to provide highly focused performances.

Sub-task 1.1.1: Collection of the Consumer Needs

From the knowledge-base collected for this case-study (§4.1), consumers want from a hand-wash the following consumer needs (ψ_i):

- High foam-ability;
- Non irritability of the skin;
- Wetting of the skin;

- Dissolution of the dirt;
- Suspension of the dirt;
- Spread-ability;
- Good stability;
- Anti-bacterial performances;
- Pleasant color;
- Pleasant odor;
- Pleasant skin feeling;
- Low toxicity;
- High safety.

As it has been underlined in §6.2.1, the consumer need for cleaning performances is here divided in its three main constituent (wetting of the substrate/skin, dissolution of the dirt, suspension of the dirt) for a better identification of the relative target properties.

Table 6.21 Distinction between main and secondary consumer
needs for a hand-wash in the emulsified form

Consumer Needs (ψ_i)	Main Consumer Needs (ψ_M)	Secondary Consumer Needs (ψ_S)
High foam-ability	✓	
Non irritability of the skin	✓	
Wetting of the substrate	✓	
Dissolution of the dirt	✓	
Suspension of the dirt	✓	
Spread-ability		✓
Good stability		✓
Anti-bacterial performances		✓
Pleasant color		✓
Pleasant odor		✓
Pleasant skin feeling		✓
Low toxicity	✓	
High safety	✓	

Sub-task 1.1.2: Definition of the Main and the Secondary Consumer Needs

According to the knowledge-base, the main consumer needs (ψ_m) are defined, as opposed to the secondary consumer needs (ψ_s). This classification is shown in Table 6.21.

Sub-task 1.2.1: Collection of the Target Properties I – Not Modelled

In relation to this specific case study, eight classes of ingredients (ξ_i) are required for the satisfaction of the consumer needs collected in the previous task. They are all listed in Table 6.22.

Table 6.22 Collection of the classes of ingredients and distinction between active ingredients and additives for a hand-wash in the emulsified form

Consumer Needs (ψ_i)	Class of Ingredients (ξ_i)	Active Ingredients (ξ_{AI}) or Additives (ξ_{ADD})
High foam-ability	Ionic Surfactant	ξ_{AI}
Non irritability of the skin		
Wetting of the substrate		
Dissolution of the dirt	Non-Ionic Surfactant	ξ_{AI}
Suspension of the dirt		
Spread-ability	-	-
Good stability	Co-surfactant	ξ_{ADD}
	Preservative	ξ_{ADD}
Anti-bacterial performances	Anti-bacterial Agent	ξ_{ADD}
Pleasant color	Colorant	ξ_{ADD}
Pleasant odor	Aroma	ξ_{ADD}
Pleasant skin feeling	Emollient	ξ_{ADD}
Low toxicity	-	-
High safety	-	-

The choice of two surfactants is because ionic surfactants are recognized to produce higher amounts of foam, as opposed to non-ionic surfactants, while non-ionic surfactants are usually milder on the skin. Therefore, both are selected.

Sub-task 1.2.2: Definition of the Active Ingredients and of the Additives

Active ingredients and additives are defined by following the distinction between main (ψ_m) and secondary (ψ_s) consumer needs, respectively. According to such a classification, reported in Table 6.21 in relation to this case study, ionic and non-ionic surfactants are defined as the active ingredients of this product, while an anti-bacterial agent, an aroma, a colorant, a preservative and a softener are defined as additives. Table 6.22 illustrates such an identification.

Table 6.23 Collection and classification of the target properties,
for a hand-wash in the emulsified form

Consumer Needs (ψ_i)	Target Properties (ζ_i)	Categories of Target Properties (ζ_i)
High foam-ability	Surface Tension	ζ_1
	Critical Micelle Concentration	ζ_1
Non irritability of the skin	Solubility Parameters	ζ_4
	pH	ζ_4
Wetting of the substrate	Surface Tension	ζ_1
Dissolution of the dirt	Solubility Parameters	ζ_1
Suspension of the dirt	Hydrophilic-Lipophilic Balance	ζ_1
Spread-ability	Molar Volume	ζ_4
	Viscosity	ζ_4
	Cloud Point	ζ_1
Good stability	Krafft Temperature	ζ_1
	Hydrophilic-Lipophilic Deviation	ζ_4
Anti-bacterial performances	-	-
Pleasant color	-	-
Pleasant odor	-	-
Pleasant skin feeling	-	-
Low toxicity	Toxicity Parameter	ζ_3
High safety	Flash point	ζ_3

Sub-task 1.3.1: Collection of the Target Properties II – Modelled

Besides a list of necessary classes of chemicals, also a set of target thermo-physical properties is collected from the consumer needs identified in task 1.1. Thirteen target properties (that is, target properties II, ζ_i) are necessary for this case study and they are collected in Table 6.23.

Table 6.24 Collection of the target values and of the boundaries of acceptance of the target properties, for a tank-cleaning detergent in the emulsified form

Target Properties (ζ_i)	Target Value ($\zeta_{i,T}$)	Boundaries of Acceptance $\zeta_{i,L}$, $\zeta_{i,U}$	Unit of Measure
Surface Tension	-	$\sigma < 25$	mN/m
Critical Micelle Concentration	-	CMC < 0.01	mol/l
	-	$\delta_D < 20.4$ and $\delta_D > 24.4$	MPa$^{1/2}$
Solubility Parameters	-	$\delta_P < 7.8$ and $\delta_P > 11.8$	MPa$^{1/2}$
	-	$\delta_H < 9.9$ and $\delta_H > 13.9$	MPa$^{1/2}$
pH	pH = 5.5	4.5 < pH < 8	[-]
Surface tension	-	$\sigma < 55$	mN/m
Solubility Parameters	$\delta_T = 17.7$	$14.7 < \delta_T < 20.7$	MPa$^{1/2}$
Hydrophilic-Lipophilic Balance	-	HLB > 12	[-]
Molar Volume	-	$30 < V_m < 150$	l/kmol
Viscosity	-	$5 < \mu < 2500$	cP
Cloud Point	-	CP > 70	°C
Krafft temperature	-	$T_K < 20$	°C
Hydrophilic-Lipophilic Deviation	-	HLD $\neq$ 0	[-]
Toxicity Parameter	-	$LC_{50} > 3.16$	mol/m^3
Flash Point	-	$T_f > 70$	°C

Sub-task 1.3.2: Classification of the Target Properties

Among the thirteen target properties (ζ_i) identified in the previous sub-task 1.3.1, seven are relative to the active ingredients (ζ_1), two are relative to the overall product (ζ_4), and two are relative to all the formulation ingredients (ζ_3). None of the target properties collected is classified as relative to the additives (ζ_2). The overall classification is given in Table 6.3.

The same comment on the surface tension reported in §6.2.1 is valid here, as the surface tension of the overall product is determined by the pure component property of the surfactants. In addition, as the need for high foam-ability is related to the presence of an ionic surfactants, the relative target properties are requirement only of the ionic surfactant. Similarly, the target properties identified for the cleaning performances are relative to the non-ionic surfactant, among the active ingredients.

Sub-task 1.3.3: Identification of the Target Values and of the Boundaries of Acceptance

For each of the target properties retrieve in sub-task 1.3.1, a target value and/or boundaries of acceptance have to be set, for ingredient screening and selection reasons. Table 6.14 lists all the numerical constraints defined for the collected target properties.

It needs to be noticed that the boundaries of acceptance of the Hansen solubility parameters are defined through the Hansen solubility parameters of the proteins of the skin. Moreover, two different boundaries of acceptance are set for the surface tension, but as underlined in sub-task §1.3.2, they refer to different active ingredients.

6.3.2 Stage 2: Model-based Stage

Model-based techniques are adopted in this stage for generating a candidate emulsified formulated product to be further verified in the experiment-based stage.

Sub-task 2.1.1: Screening of the Candidate Active Ingredients

Two databases have to be retrieved in this sub-task, according to the information collected in sub-task 1.2.2: the ionic surfactants database, and the non-ionic surfactants database. In relation to the ionic surfactants, constraints on the surface tension ($\sigma < 25$ mN/m), on the critical micelle concentration and on the Krafft temperature are applied, while for the non-ionic surfactants, constraints on the surface tension ($\sigma < 55$ mN/m), on the Hildebrand solubility parameter, on the hydrophilic-lipophilic balance and on the cloud points are considered. For both the active ingredients, in addition, the constraints on the toxicity parameter and on the flash point are also set. A short-list of candidate ionic and non-ionic surfactants is then generated.

Sub-task 2.1.2: Selection of the Best Active Ingredients

Also in relation to this case study, the effectiveness is chosen as the selection criteria for the active ingredient. In relation to the ionic surfactant, the lower the surface tension and the critical micelle concentration, the higher the performances; while for the non-ionic surfactants, the effectiveness is identified with the target values set for the Hildebrand solubility parameter. In both cases, when two or more candidate ingredients have comparable effectiveness, the cheapest solution is selected.

As a result, sodium dodecyl sulfate and tween 60 are selected as the best ionic and non-ionic surfactants, respectively. Their chemical structures are given in Figure 6.5.
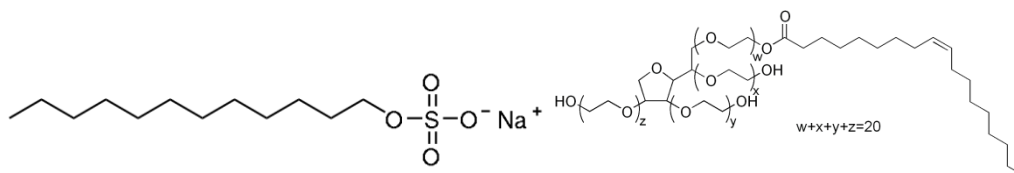
Figure 6.5 Chemical structures of sodium dodecyl sulphate and tween 60,
chosen as active ingredients for a hand-wash in the emulsified form

Sub-task 2.1.3: Collection of the Active Ingredients Properties

The necessary thermo-physical properties of the two active ingredients are collected here,
and calculated by means of pure component property models, if necessary. They are given
in Table 6.25.

Table 6.25 Candidate active ingredients with thermo-physical properties,
for a hand-wash in the emulsified form

Comm. Name	CMC [mol/L]	T_k/CP [°C]	σ [mN/m]	HLB [-]	MW [g/mol]	ρ [g/ml]	μ [cP]	T_f [°C]	LC_{50} [mol/m³]	Cost [$/kg]
Sodium Dodecyl Sulfate	0.004	16.0	24.8	40	288.37	1.01	197	>100	3.51	163
Tween 60	0.011	94	31	13.4	438.73	1.04	450	>100	5.04	203

Sub-task 2.2.1: Retrieval of the Databases

Hand-wash products are usually oil-in-water emulsions, therefore it is necessary to re-
trieve one database for the aqueous solvent phase, and one for the organic solvent phase.
In fact, similarly to the case study of §6.2, two surfactants have been selected as active
ingredients, therefore there is no need for selecting the ingredients of the surfactant sys-
tem in the emulsified solvent mixture.

In relation to the aqueous solvent phase, instead, water is chosen as the product is directly
applied on the skin, and the non-irritability of the skin is one of the requirements collected
in Table 6.21.

In relation to the organic solvent phase, finally, vegetable oils are often used in the cos-
metic products for their moisturizing properties as well as their pleasant aroma. The rel-
ative database is then retrieved for this case study.

Sub-task 2.2.2: Choice of the Property Models

As in the previous case studies, the models collected in Table 3.4 are applied in the
EMUD algorithm.

Sub-task 2.2.3: Application of the EMUD Algorithm

The constraints that have been set in the problem definition stage are employed, in relation to the databases selected in sub-task 2.2.1 and the property models chosen in sub-task 2.2.2. Here is a list of constraints on the pure and mixture properties:

- Pure component properties

  Aqueous solvent phase: $LC_{50} > 3.16$ mol/m^3

                                         $T_f > 70°C$

  Organic solvent phase: $LC_{50} > 3.16$ mol/m^3

                                         $T_f > 70°C$

- Mixture properties with linear mixing rule

  $$30 < V_m < 150 \text{ l/kmol}$$
  $$\delta_D < 20.4 \text{ and } \delta_D < 20.4 \text{ MPa}^{1/2}$$
  $$\delta_P < 7.8 \text{ and } \delta_P < 11.8 \text{ MPa}^{1/2}$$
  $$\delta_H < 9.9 \text{ and } \delta_H < 13.9 \text{ MPa}^{1/2}$$
  $$C_{surf} \geq 10 \cdot CMC_{surf}$$

- Mixture properties with non-linear mixing rule

  $$5 < \mu < 2500 \text{ cP}$$

- Stability check

  $$HLD \neq 0$$

Under the list of constraints given above constraints, the optimal emulsified solvent mixture generated from the EMUD algorithm is as follows:

- Water (aqueous solvent phase): 57 % in weight;
- Jojoba oil (organic solvent phase): 27% in weight;
- Sodium dodecyl sulfate (ionic surfactant): 8% in weight;
- Tween 60 (non-ionic surfactant): 8% in weight.

The thermo-physical properties of jojoba oil are given in Table 6.26.

Table 6.26 Thermo-physical properties of jojoba oil

Comm. Name	δ_D [MPa$^{1/2}$]	δ_P [MPa$^{1/2}$]	δ_H [MPa$^{1/2}$]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [°C]	Toxicity Parameter [mol/m^3]	Cost [$/kg]
Jojoba Oil	16.0	2.8	6.2	295	0.88	35	>100	5.32	182

Sub-task 2.2.4: Collection of the Emulsified Solvent Mixture Properties

Rigorous models are used in this sub-task to verify that solvent mixture properties satisfy the constraints previously set. The results are summarized below:

- Dynamic viscosity: 7.03 cP (lower boundary: 5; upper boundary: 2500 cS);
- Molar volume: 30.2. l/kmol (lower boundary: 30; upper boundary: 150 l/kmol);

- Hansen solubility parameters: dispersion: 27.4 MPa$^{1/2}$

                                           polar: 14.5 MPa$^{1/2}$

                                           hydrogen-bonding 16.2 MPa$^{1/2}$
- Hydrophilic-lipophilic deviation: -3.1 (lower boundary: 0).

These results satisfy the constraints set for the solvent mixture in sub-task 2.2.3.

Sub-task 2.3.1: Screening of the Candidate Additives

Table 6.22 highlight the need for six additives to be included in the emulsified formulated hand-wash, to fulfill the secondary consumer needs: an anti-bacterial agent, an aroma, a colorant, a co-surfactant, an emollient, and a preservative. However, the Food and Drug Administration (FDA) has recently expressed concerns over anti-microbial agents, as they are currently under investigation for potential carcinogen effects. For this reason, the anti-microbial agent is excluded from the list of the necessary additives.

Four additives databases are then retrieved: the aroma database, the colorant database, the co-surfactant database, the emollient database, and the preservative database. Constraints on the pure component properties in terms of solubility parameters, flash point and toxicity parameters are set:

- Solubility parameter $14.4 < \delta_T < 20.4$ MPa$^{1/2}$ (organic phase)

                                               $27.1 < \delta_T < 33.1$ MPa$^{1/2}$ (aqueous phase)
- Flash point $T_f > 70°C$
- Toxicity parameter $LC_{50} > 3.16$ mol/m^3

Four short-lists of candidate additives are then generated accordingly.

Sub-task 2.3.2: Selection of the Best Additives

The criteria chosen for the selection of the best candidate additives is the cost. The cheapest ingredients satisfying the constraints set in sub-task 2.3.1 is selected. The four candidate additives are then:

- Aroma: Orange sweet oil
- Colorant: Orange colorant
- Co-surfactant: Propylene glycol
- Emollient: Polyquaterium-7
- Preservative: Sodium benzoate

Figure 6.6 illustrates the chemical structures only of propylene glycol and sodium benzoate, as the other additives are mixtures and not pure ingredients.
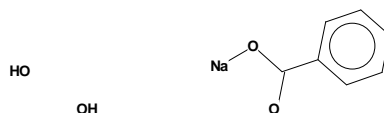
Figure 6.6 Chemical structure of propylene glycol (left) and sodium benzoate (right), chosen as additives for a hand-wash in the emulsified form

Sub-task 2.3.3: Collection of the Additives Properties

In this sub-task, the target properties of the selected additives are retrieved from the databases, or estimated with property models. The results are collected in Table 6.27.

Table 6.27 Candidate additives with thermo-physical properties, for a tank-cleaning detergent in the emulsified form

Commercial Name	Qualitative Solubility [-]	MW [g/mol]	Density [g/ml]	Liquid Viscosity [cP]	Flash Point [˚C]	Toxicity Parameter [mol/m³]	Cost [$/kg]
Orange Sweet Oil	organic	136.23	0.84	1.32	92	3.55	92
Orange Colorant	organic	-	0.97	1.12	>100	7.18	13
Propylene Glycol	water	76.09	1.04	52.4	96	6.12	63
Polyquaternium-7	water	63.87	1.02	-	>100	3.71	51
Sodium Benzoate	water	144.10	1.50	12	>100	4.02	50

Sub-task 2.4.1: Calculation of the Product Properties

In this sub-task, rigorous mixture property models are applied for verification of the target mixture properties. In fact, the addition of active ingredients and additives to the emulsified solvent mixture altered its properties, therefore it is necessary to verify that the product properties still match the constraints.

The result of the calculation is summarized below:

- Dynamic viscosity: 14.11 cP (lower boundary: 5; upper boundary: 2500 cS);
- Molar volume: 32.7. l/kmol (lower boundary: 30; upper boundary: 150 l/kmol);

The results obtained are in agreement with the constraints previously set.

Sub-task 2.4.2: Evaluation of the Product Stability

In this last sub-task, the HLD value of the product is calculated, to check if the addition of active ingredients and additives has influenced the stability of the product as an emulsion. The new calculated HLD value is -3.8, showing a decrease compared to the solvent mixture, indicating that the stability of the product has not been compromised.

The final candidate emulsified formulated product generated in the model-based design stage is reported in Table 6.28.

Table 6.28 Candidate formulation for a hand-wash in the emulsified form

	Class of Ingredient	Commercial Name	Weight Percentage
Active Ingredients	Ionic Surfactant	Sodium dodecyl sulfate	7.5
	Non-Ionic Surfactant	Tween 60	7.5
Solvent Mixture	Aqueous solvent	Water	54
	Organic solvent	Jojoba oil	25
Additives	Aroma	Orange Sweet Oil	1.5
	Colorant	Orange Colorant	1
	Co-surfactant	Propylene Glycol	2.5
	Emollient	Polyquaternium-7	0.5
	Preservative	Sodium Benzoate	0.5

6.3.3 Stage 3: Experiment-based Stage

In relation to this specific case study, the overall integrated methodology has been applied, that is, the experiment-based stage has been planned and the relative experiments have been performed.

Sub-task 3.1.1: Collection of the Product Properties to be validated

In this sub-task, the product properties to be measured for validation are collected. They consist of those target properties identified in sub-task 1.3.1. A list, comprehensive of the property classification, is given below:

- Surface tension (ζ_1);
- Critical micelle concentration (ζ_1);
- Solubility parameters (ζ_4);
- pH (ζ_4);
- Surface tension (ζ_1);
- Solubility parameters (ζ_1);
- Hydrophilic-lipophilic balance (ζ_1);
- Molar volume (ζ_4);
- Dynamic viscosity (ζ_4);
- Cloud point (ζ_1);

- Krafft temperature (ζ_1);
- Hydrophilic-lipophilic deviation (ζ_4);
- Toxicity parameter (ζ_3);
- Flash point (ζ_3).

Sub-task 3.1.2: Collection of the Product Performances to be validated

In this sub-task, the product performances to be assessed for validation are collected. They consist of those target properties identified in sub-task 1.2.1. The list of these product performances is given below:

- High foam-ability;
- Non irritability of the skin;
- Cleaning performances;
- Spread-ability;
- Good stability;
- Anti-bacterial performances;
- Pleasant color;
- Pleasant odor;
- Pleasant skin feeling:
- Low toxicity;
- High safety.

Sub-task 3.1.3: Generation of a List of Experiments

Based on the product properties and performances collected in the previous sub-tasks, a list of experiments for validation is generated. Table 6.29 list the proposed experiments, in relation to each of the target properties and performances to be validated.

Sub-task 3.2.1: Collection of the possible Non-validations

The reasons for which the properties and performances assessed in this stage might not match with the predictions are collected in this sub-task. However, as in §6.1 and §6.2, the results are not generated as they can be better illustrated by means of the list of action.

Table 6.29 List of experiments for a hand-wash in the emulsified form

Consumer Needs (ψ_i)	Product Property Validation Experiment	Product Performance Validation Experiment
High foam-ability	-	Panel test for the foam-ability
Non irritability of the skin	Measurement of the pH	Panel test for the irritability of the skin
Wetting of the substrate	Measurement of the surface tension	-
Dissolution of the dirt	-	-
Suspension of the dirt	-	-
Spread-ability	Measurement of the molar volume / Measurement of the viscosity	Panel test for the spread-ability
Good stability	Measurement of the phase inversion temperature	Solubility test / Solvent mixture stability test / Shelf-life test
Anti-bacterial performances	-	-
Pleasant color	-	Panel test for the color
Pleasant odor	-	Panel test for the odor
Pleasant skin feeling	-	Panel test for the skin feeling
Low toxicity	-	-
High safety	-	-

Sub-task 3.2.2: Generation of a List of Actions

In this sub-task, a list of action corresponding to every validation experiment defined in sub-task 3.1.3 is generated. Table 3.20 provides the list of action built here.

Table 6.30 List of actions for a hand-wash in the emulsified form

Validation Experiment	Action 1	Action 2
Measurement of pH	Add an adequate additive	Change the additive
Measurement of σ	Change the solvent mixture	Change the active ingredient
Measurement of V_m	Add an adequate additive	Change the solvent mixture
Measurement of μ	Change the droplet size distribution	Add an adequate additive
Measurement of the PIT	Add an adequate additive	Change the solvent mixture
Panel test for the foam-ability	Change the concentration of the active ingredient	Change the active ingredient
Panel test for the irritability of the skin	Add an adequate additive	Change the additive
Panel test for the spread-ability	Change the solvent mixture	-
Solubility test	Change the solvent mixture	-
Solvent mixture stability test	Change the solvent mixture	-
Shelf-life test	Change the solvent mixture	-
Panel test for the color	Add an adequate additive	Change the additive
Panel test for the odor	Add an adequate additive	Change the additive
Panel test for the skin feeling	Add an adequate additive	Change the solvent mixture

Sub-task 3.3.1 Experimental Availability Check

All the necessary chemicals listed in Table 6.28 are available. In terms of experimental setup, instead, a series of equipment and experimental procedures are chosen, on the bases of the setup availability. Moreover, the experiments are listed accordingly to their difficulty and time length: from the most simple and/or fast to the most difficult and/or time consuming.

1. Solubility test: the solubility of the active ingredients and of the additives in the respective solvent phases is assessed. As the aim of this test is not to identify the solubility limit, but to check if the designed formulation was stable at the designed concentration, these tests are performed by mixing and observing if phase stability occurs;

2. Solvent mixture stability test: the solvent mixture is manufactured at the designed optimal composition. The sample is mixed for 30 minutes at 1500 rpm in a mechanical stirrer (IKA T25 digital ultra-torrax) and it is then transferred in an equilibrium glass with an external jacket for temperature control. The temperature is

changed at the speed of 1°C per minute, and the conductivity is monitored in continuous (with a waterproof hand-held conductivity TDS meter, WPA CMD410). The temperature at which a sharp discontinuity in the conductivity is observed is the phase inversion temperature, and it is considered the stability boundary, with respect to the temperature;

3. A product prototype is manufacture by first homogenizing the two liquid phases, by mixing for 30 minutes at 1000 rpm with the same mechanical stirrer used for the previous experiment; then the dispersed phase (that is, the organic phase) is slowly poured in the aqueous phase, and the formulation is mixed for 30 more minutes at 1500 rpm;

4. Measurement of the pH: indicator strips (Merck) are used to measure the pH of the product prototype;

5. Measurement of the molar volume: a known volume of the prototype is weighted and then the molar volume is obtained by multiplying the ratio between the volume and the mass per the average molecular weight of the product;

6. Measurement of the dynamic viscosity: a sample of the prototype is poured into a capillary viscometer and the capillary rise speed is measured; through the characteristic equation of the equipment, the dynamic viscosity is calculated;

7. Measurement of the surface tension: a drop of the product prototype is poured on a glass surface and the contact angle is calculated by taking a high-resolution picture and using a dedicated software; then the equation of the capillary action is used to estimate the surface tension;

8. Measurement of the phase inversion temperature: the same procedure of the experiment number 2 is applied to a sample of the product prototype;

9. Panel test for spread-ability: a prototype sample is poured into a commercial hand-wash dispenser and the functionality of the device is tested;

10. Panel test for cosmetic properties: the foam-ability, the non-irritancy to the skin, the pleasant odor, the pleasant color and the skin feeling are evaluated by applying a sample of the prototype on the skin;

11. Shelf-life test: no adequate equipment has been retrieved for performing this experiments.

Sub-task 3.3.2 Experiments

The experiments listed in sub-task 3.3.1 were performed in this sub-task to verify and amend the candidate emulsified formulated product. The results are stored for comparison with the predictions.

Sub-task 3.4.1 Comparison between Model-based Calculation and Experimental Results

In this sub-task, the results of the experiments performed in the previous sub-task are compared with the model-based predictions and with the constraints set in the problem definition stage.

1. Solubility test: the orange sweet oil and the orange colorant has been found to be soluble in the jojoba oil up to 50% in weight; higher concentrations have not been verified as this was out of the significance of these experiments; propylene glycol and sodium benzoate are found to be soluble in deionized water up to 50% in weight; higher concentrations have not been verified as above; Polyquaternium-7 has been found to be soluble up to 20% in weight;

2. Solvent mixture stability test: the observed phase inversion temperature of the emulsified solvent mixture is 65˚C, compared with a prediction (via the HLD approach) of 68 ˚C;

3. The product prototype is manufactured;

4. The measured pH of the prototype is 8, corresponding to the upper boundary set in the problem definition stage;

5. Measurement of the molar volume: the measured molar volume is 31.1 l/kmol, compared to a prediction of 32.7 l/kmol and lower and upper boundaries set to 30 and 150 l/kmol, respectively;

6. Measurement of the dynamic viscosity: the measured dynamic viscosity is 18 cP, compared to a prediction of 14.11 cP and lower and upper boundaries set to 5 and 25000 cP, respectively;

7. Measurement of the surface tension: the measured surface tension is 24 mN/m, compared to a prediction of 21 mN/m and the lower boundary set to 25 mN/m;

8. Measurement of the phase inversion temperature: the observed phase inversion temperature is 65˚C, compared to a prediction of 73˚C;

9. Panel test for spread-ability: the prototype sample flows through the commercial hand-wash dispenser;

10. Panel test for cosmetic properties: the foam-ability is considered sufficient; the skin feeling is pleasant (no irritancy has been detected, even 48 hours after the application), but a higher viscosity is preferred; the odor is considered very pleasant, and persistent even a few hours after the application; the color is considered also pleasant;

11. Shelf-life test: the test has not been performed, however a few days after the manufacture of the product prototype, signs of instability are observed (tendency to creaming).

Sub-task 3.4.2 Generation of a List of Refinements

The main output that are obtained from sub-task 3.4.1 are summarized here:

1. The solubility of the active ingredients and of the additives in the respective solvent phases has been successfully verified;
2. The stability of the solvent mixture at the designed composition has been successfully verified, and the prediction has been found in good agreement with the observation;
3. -;
4. The pH has been found to correspond to the higher boundary of acceptance set; actions must be taken;
5. The molar volume has been found to be successfully into the boundaries of acceptance set, and the prediction has been found in good agreement with the measurement;
6. The dynamic viscosity has been found to be successfully into the boundaries of acceptance set, and the prediction has been found in good agreement with the measurement;
7. The surface tension has been found to be just into the boundaries of acceptance set, and the prediction has been found quite in agreement with the measurement;
8. The prediction of the phase inversion temperature has been found to be quite in agreement with the observed value;
9. The spread-ability through a commercial dispenser has been successfully verified;
10. The foam-ability, the non-irritability of the skin, the odor, and the color has been successfully verified; the product viscosity is considered insufficient, from the point of view of the skin feeling; actions must be taken;
11. The stability of the product prototype has not been verified; however the tendency to creaming observed a few days after the product manufacture suggests a non-acceptable product stability; actions must be taken.

According to the list above, actions to lower the product pH, to increase the product viscosity and to improve the product stability must be taken. The list of action generated in sub-task 3.2.2 is then used to generate a list of refinements. The final results are presented below:

*   The addition of a mild acidic buffer is necessary to reduce the pH; 0.5% in weigh of citric acid (water-soluble) is added to the emulsified formulation;
*   The addition of a thickener is necessary to increase the viscosity; 1% in weight of polyethylene glycol (water-soluble) is added to the emulsified formulation;
*   The reduction of the average droplet size and the achievement of a narrower droplet size distribution are necessary to increase the stability; a different stirring equipment, characterized by higher stirring speed is recommended.

From the point of view of the product formulation, therefore, Table 6.31 presents the refinement composition of a hand-wash in the emulsified form, obtained by a first model-based design technique, further refined through an experiment-based verification stage.

Table 6.31 Refined formulation for a hand-wash in the emulsified form

	Class of Ingredient	Commercial Name	Weight Percentage
Active Ingredients	Ionic Surfactant	Sodium dodecyl sulfate	7.4
	Non-Ionic Surfactant	Tween 60	7.4
Solvent Mixture	Aqueous solvent	Water	53.1
	Organic solvent	Jojoba oil	24.6
Additives	Aroma	Orange Sweet Oil	1.5
	Colorant	Orange Colorant	1
	Co-surfactant	Propylene Glycol	2.5
	Emollient	Polyquaternium-7	0.5
	Preservative	Sodium Benzoate	0.5
	Acidic buffer	Citric Acid	0.5
	Thickener	Polyethylene Glycol	1

7

CONCLUSIONS

In this work, a special branch of chemical product design has been investigated. Consumer oriented chemical based products have been considered, with focus on emulsified formulated products, where surfactants play a major role for the stability of the product in the emulsified physical form. This work proposes an integrated, systematic methodology, where computer-aided tools are used at a first level, for screening among a wide range of alternatives, and then experimental techniques are applied on a short number of candidate formulations for validation and refinement.

In this chapter, first the achievements of this work are listed (§7.1), and the a few perspective on future work are given (§7.2).

7.1 Achievements

In Chapter 2, a list of issues and needs to be addressed was presented. It consists of:

- Development of a robust knowledge base;
- Development of structured ingredients databases;
- Development of property models for surfactants;
- Development of an algorithm for the design of emulsified solvent mixtures;
- Development of an integrated methodology for the design of emulsified formulated products;
- Development of a systematic framework, based on the above mentioned methodology, including the necessary knowledge base, databases, property models and algorithms;
- Development of case studies to test and validate the proposed methodology.

All the requirements above have been addressed and achieved in this work.

All the information that are necessary for supporting the decisions and choices performed during the design of emulsified formulated products have been collected and stored in a systematic knowledge base. The knowledge base is organized in layers, so that the required information can be retrieved and used in an easy and functional way, when needed. The presence of the knowledge base allows a reliable solution of the problem definition stage, generating a list of useful information (necessary classes of ingredients and target thermo-physical properties) that are used as the input of the model-based and experiment-based stage.

Databases of chemicals have been retrieved, improved and developed, to assist the ingredient selection and design in the model-based stage. The databases are structured so that the ingredients are divided for their role in the formulation (active ingredients, additives, solvents and surfactants) and for their main function (aromas, colorants, *etc.*). In addition, all the thermo-physical properties that are required in the design stage have been retrieved or calculated for all the available chemicals.

Accurate property models for the target thermo-physical pure component as well as mixture properties that are necessary in the design of emulsified formulated products have been retrieved from the literature, and adapted for the estimation of the needed target properties. If these property models were not available, and/or their accuracy was not satisfactory, new property models have been developed.

Group-contribution property models based on the Marrero and Gani method have been developed for the prediction of the cloud point and of the critical micelle concentration of non-ionic surfactants. The statistical indices have been compared with those of other property models, based on the QSPR approach, and the group-contribution models show better performances. A new thermodynamic approach for the assessment of the stability of emulsified solvent mixtures, in relation to temperature and concentration, has also been proposed. However, the lack of experimental data represents a problem for regressing new parameters of a thermodynamic model, such as UNIFAC; therefore the application of this approach in the design procedure is limited. It can be used for verification purposes, as long as the necessary experimental data are available.

A new algorithm for the design of emulsified solvent mixtures (EMUD) has been developed. The EMUD algorithm is a multi-level algorithm, based on the reverse approach: given the desired set of target property values, the algorithm identifies the emulsified solvent mixtures (that is, an aqueous solvent, an organic solvent, and a surfactant system) matching the constraints. The constraints are applied at four different levels: first, constraints on pure component properties are applied, and the candidate ingredients are screened on these basis; then, constraints on mixture properties described by the linear mixing rule are employed and the compositions minimizing the cost are identified. Then, constraints on mixture properties described by non-linear mixing rules are applied, and

finally, the stability of the candidate solvent mixture in the emulsified form is assessed by means of a correlation. The integration of the EMUD algorithm with another algorithm previously developed for the design of binary homogeneous solvent mixture in the liquid phase (MIXD), is also discusses.

An integrated methodology for emulsified formulated product design has been proposed. It consists of three stages: a problem definition stage, a model-based design stage, and an experiment-based verification stage. In the first stage, thanks to the knowledge base, the consumer needs are collected and translated into a set of necessary classes of ingredients and of target thermo-physical properties, with target values and/or boundaries of acceptance. The second stage applies computer-aided techniques, structured databases, property models and the EMUD algorithm to generate a candidate formulation to be verified by means of experiments, from the information obtained in the previous stage. The third step, finally, adopt experimental techniques for verification of the candidate product. The properties and performances of the proposed formulation are assessed and, in case there is no match between any of the measured property or performance and its prediction, a list of action generated with the knowledge base is applied for refining the proposed formulation.

A systematic framework for the design and analysis of emulsified formulated products has been included as a new template to an existing software: the virtual Product-Process Design Laboratory, thus extending its application range. The framework is based on the problem definition and model-based design stage of the methodology for emulsified formulated product design, and it collects the knowledge-base, the databases, the models and the algorithm developed in this work.

Finally, the systematic methodology for emulsified formulated product design, and the corresponding framework, have been tested on a number of case studies. Three consumer oriented chemical based products in the emulsified form have been considered: a UV sunscreen, a tank-cleaning detergent and a hand-wash. For the first two case-studies, only the first two stages and part of the experiment-based stage have been applied; that is, the experimental work has been planned, the list of actions has been generated, but experiments have not been performed. In relation to the hand-wash, instead, the whole methodology has been applied and a final prototype has also been manufactured. The results confirm that the integrated methodology developed in this work can be successfully employed for the screening of thousands of alternatives, and for the generation of a short-list of promising candidates. Some improvements have been proposed in relation to some property models that are to be improved for a more reliable estimation of the product properties.

7.2 Future Work

One of the main challenges of product design identified in Chapter 2 is related to the problem definition. This work has tried to systematize this task, developing a robust knowledge base for the collection of the consumer needs and their translation into necessary categories of ingredients and target properties with target values and boundaries of acceptance. However, the knowledge base can be further improved by involving experts in the area to share their knowledge and experience, and by adding more information relative to a wider variety of chemical products.

In relation to some of the target properties, such as, cosmetic and sensorial properties, the lack of quantitative models affects the applicability of computer-aided tools for the design and selection of the relative ingredients. Additional efforts should be focused in relating these factors with thermo-physical properties, so that the relative ingredients can be designed on quantitative bases, and therefore time and resources during the experimental validation can be spared, thanks to their inclusion in the model-based design stage.

The relation between the product performances, such as, the sun protection factor for a UV sunscreen, and the concentration of the active ingredients in the solvent mixture is another area on which future research could focus. In this work, the information collected and stored in the knowledge base, from literature and patents, are used, limiting the potential for product innovation.

The scarce availability of experimental data relative to the phase equilibria of systems containing surfactants, limits the possibility of the development of a thermodynamic approach for the assessment of the stability of a product in the emulsified form. Such an approach is considered a valid alternative to the heuristics and correlations that are currently the state of the art. The regression of the parameters of a truly predictive model, such as UNIFAC, on the basis of such phase equilibria data, is regarded as a relevant advance in understanding and modeling of emulsified products.

Finally, when designing structured products as emulsified formulation, it is recommended that the product and the manufacturing process are designed simultaneously. Structured products, in fact, achieve their properties through an internal structure that is determined by the interaction of its component and the manufacturing process. The recent advances achieved by the process systems engineering community in the area of the super-structure optimization can be applied to the solution of such a complex problem, as long as the necessary product and process models are available.

.

APPENDICES

Appendix A: Modeling of Cloud Point of Non-ionic Surfactants with Group-Contribution Model

Table A.1 List of the original data-set of experimental cloud point (in °C) of
nonionic surfactants (1% weight percentage). Distinction in classes
is based only on the molecular structure, each of which is described.
Compounds highlighted in grey have been excluded from the parameter regression step.

Code	CP_{exp} [°C]	Code	CP_{exp} [°C]	Code	CP_{exp} [°C]
Linear alkyl ethoxylates					
CnEm		**$C_nH_{2n+1}O(C_2H_4O)_mH$**			
C4E1	44.5	C9E6	75	C12E10	95.5
C5E2	36	C10E4	19.7	C12E11	100.3
C6E2	0	C10E5	41.6	C13E5	27
C6E3	40.5	C10E6	60.3	C13E6	42
C6E4	63.8	C10E7	75	C13E8	72.5
C6E5	75	C10E8	84.5	C14E5	20
C6E6	83	C10E10	95	C14E6	42.3
C7E3	27.6	C11E4	10.5	C14E7	57.6
C8E3	7	C11E5	37	C14E8	70.5
C8E4	38.5	C11E6	57.5	C15E6	37.5
C8E5	58.6	C11E8	82	C15E8	66
C8E6	72.5	C12E4	6	C16E6	35.5
C8E8	96	C12E5	28.9	C16E7	54
C8E9	100	C12E6	51	C16E8	65
C8E12	106	C12E7	64.7	C16E9	75
C9E4	32	C12E8	77.9	C16E10	66
C9E5	55	C12E9	87.8	C16E12	92

Branched alkyl ethoxylates					
ICnEm		$(C_{(n-2)/2}H_{n-1})_2CHCH_2O(C_2H_4O)_mH$		(n = 6, 10)	
		$(C_{(n-1)/2}H_n)_2CHO(C_2H_4O)_mH$		(n=13)	
TCnEm		$(C_{(n-1)/3}H_{(2n+1)/3})_3CO(C_2H_4O)_mH$			
IC6E6	78	TC10E7	22	TC13E9	34
IC10E6	27	IC13E9	35	TC16E12	48
Phenyl alkyl ethoxylates					
CnPEm		$C_nH_{2n+1}C_6H_4O(C_2H_4O)_mH$			
TC8PE9	64.3	C9PE8	34	C12PE9	33
C8PE7	22	C9PE9	56	C12PE11	50
C8PE9	54	C9PE10	75	C12PE15	90
C8PE10	75	C9PE12	87		
C8PE13	89	C9PE13	89		
Alkyl polyoxyethylene-polyoxypropylene copolymers					
CnEmPk		$C_nH_{2n+1}C_6H_4O(C_2H_4O)_m(C_3H_6O)_kH$			
C12E4P5	22.1	C12E3P6	10.6	C12E5P4	29.8
Carbohydrate-derivate ethoxylates					
CnCOOEmC		$C_nH_{2n+1}COO(C_2H_4O)_mCH_3$			
CnCOOEm		$C_nH_{2n+1}COO(C_2H_4O)_mH$			
C9COOE7C	44	C9COOE12	74	C11COOE8	53
C9COOE10C	65	C11COOE6	54		
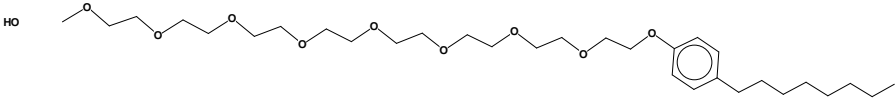
Ethoxylated amides					
CnGEm		$C_nH_{2n+1}NHCH_2COO(C_2H_4O)_mH$			
CnAEm		$C_nH_{2n+1}NHCHCH_3COO(C_2H_4O)_mH$			
CnSEm		$C_nH_{2n+1}NCH_3CH_2COO(C_2H_4O)_mH$			
C12GE2	78	C12GE4	75	C12SE3	44
C12GE3	46	C12AE3	22.5		
Alkyl branched ethoxylates					
AGM-n(3)		$C_nH_{2n+1}CH(O(C_2H_4O)_3H)_2$			
AGM-7(3)	34	AGM-11(3)	30	AGM-13(3)	29

Appendices

Table A.2 Marrero and Gani group definition and contributions after the regression
based on 72 experimental data of cloud point

1st Order Group (i)	C_i [K^2]	2nd Order Group (j)	D_j [K^2]	3rd Order Group (k)	E_k [K^2]
CH_3	6.4351e+04	AROMRINGs[1]s[4]	0	$(CH_2)_n$-$(OCH_2CH_2)_m$ (m=3, n<8) and (m=4..5, n>8)	-1.1108e+04
CH_2	-2.2149e+03			$(CH_2)_n$-$(OCH_2CH_2)_m$ (n=5)	-6.7595e+03
CH	-6.5736e+04			$(CH_2)_m$-CO-(OCH_2CH_2) (m=8)	-2.1595e+04
C	-1.4320e+05			$(CH_2)_m$- C_6H_4-(OCH_2CH_2) (m=8)	6.0521e+03
aCH	-5.8171e+03			$((CH_2)_n)_m COC_2H_4$- (n>2, m>1)	-2.4357e+04
aC- CH_2	0				
OH	-3.0249e+03				
CH_2COO	-2.7706e+03				
CH_3O	3.7198e+04				
CH_2O	8.9104e+03				
aC-O	0				
CH_2NH	0				
OCH_2CH_2OH	3.3508e+04				

Table A.3 Prediction of the cloud point of nonyl phenyl octaethylene oxide, after the introduction of new third order groups

Nonyl phenyl octaethylene glycol	Molecular structure
Molecular formula: $C_{33}H_{60}O_{10}$	

First order groups	Occurrences	Group contribution [K^2]
CH_3	1	6.4351e+04
CH_2	16	-2.2149e+03
aCH	4	-5.8171e+03
aC-CH_2	1	0
CH_2O	7	8.9104e+03
aC-O	1	0
OCH_2CHOH	1	3.3508e+04

Second order groups	Occurrences	Group contribution
AROMRINGs[1]s[4]	1	0

Third order groups	Occurrences	Group contribution
$(CH_2)_m$-C_6H_4-(OCH_2CH_2) *(m=8)*	1	6.0521e+03

$$(CP^2)_{calc} = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k = 107576.3 \ K^2$$

$$(CP)_{calc} = 327.99 \ K$$

GC-model without 3rd order groups: $(CP)_{calc} = 322.04 \ K$

$$(CP)_{exp} = 329.15 \ K$$

Appendix B: Modeling of Critical Micelle Concentration of Non-ionic Surfactants with Group-Contribution Model

Table B.1 List of the original data-set of experimental critical micelle concentration of nonionic surfactant at 25 C19, together with the molecular description of each of them. Distinction in classes is based only on the molecular structure of the surfactants. Compounds highlighted in grey have been excluded from the parameter regression step. Compounds highlighted with (**v**) have been chosen as the validation step.

Code	$-log(CMC)_{exp}$	Code	$-log(CMC)_{exp}$	Code	$-log(CMC)_{exp}$
		Linear alkyl ethoxylates (43 compounds)			
CnEm		$C_nH_{2n+1}O(C_2H_4O)_mH$			
C4E1	0.009	C10E5 (**v**)	3.100	C12E12	3.854
C4E6	0.110	C10E6	3.046	C12E14	4.260
C6E3	0.980	C10E7	3.015	C13E8	4.569
C6E4 (**v**)	1.032	C10E8	3.000	C14E6	5.000
C6E5	1.017	C10E9	2.886	C14E8	5.046
C6E6	1.164	C11E8 (**v**)	3.523	C14E9 (**v**)	5.046
C8E1 (**v**)	2.310	C12E1	4.638	C15E8 (**v**)	5.456
C8E3	2.125	C12E2	4.481	C16E6	5.780
C8E4	2.063	C12E3	4.284	C16E7	5.770
C8E5	1.959	C12E4 (**v**)	4.194	C16E8	5.921
C8E6	2.004	C12E5	4.194	C16E9	5.678
C8E9	1.886	C12E6	4.060	C16E10 (**v**)	5.699
C9E8 (**v**)	2.520	C12E7	4.086	C16E12	5.638
C10E3	3.222	C12E8	4.000		
C10E4	3.167	C12E9	4.000		
		Phenyl alkyl ethoxylates (15 compounds)			
CnPhEm		$C_nH_{2n+1}(C_6H_4)O(C_2H_4O)_mH$			
C8PhE1	4.305	C8PhE6	3.678	C8PhE30 (**v**)	3.959
C8PhE2	4.116	C8PhE7	3.602	C8PhE40	4.119
C8PhE3	4.013	C8PhE8	3.553	C9PhE2	3.377
C8PhE4	3.886	C8PhE9	3.523	C9PhE5 (**v**)	3.328
C8PhE5 (**v**)	3.824	C8PhE10	3.481	C9PhE12	3.301
		Branched alkyl ethoxylates (5 compounds)			
ICnEm		$(CH_2C_{n/2-2}H_{n-4})_2CHCH_2O(C_2H_4O)_mH$			
IC4E6	0.049	IC8E6	1.670	IC10E9	2.526
IC6E6	1.016	IC10E6 (**v**)	2.547		
		Alkanediols (5 compounds)			
C8GLYCER	2.237	**C8H17OCH2CH(OH)CH2OH**			
CnDIOL	**(n=10,12)**	$C_{n-2}H_{2n-3}CH(OH)CH_2OH$			
C10DIOL	2.638	C12DIOL	3.745		
CnDIOL	**(n=11,15)**	$C_{n-3}H_{2n-5}CH(OH)CH_2CH_2OH$			
C11DIOL (**v**)	2.638	C15DIOL	4.886		
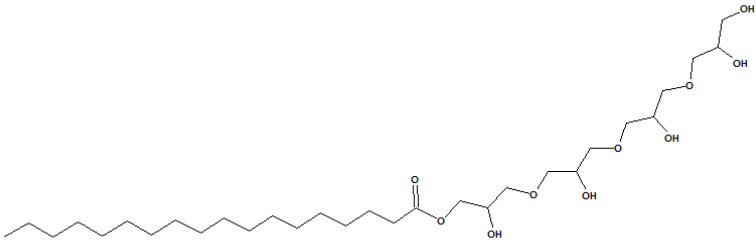
Alkyl mono and disaccharide ethers and esters (7 compounds)					
CnGLUC		$C_nH_{2n+1}O(C_6H_{11}O_5)$			
C8GLUC	*1.602*	C10GLUC **(v)**	*2.658*	C12GLUC	*3.721*
C12DELAC	*3.222*	$C_{12}H_{25}NH(C_6H_{12}O_4)O(C_6H_{11}O_5)$ *(first ring open)*			
C12MALT	*3.620*	$C_{12}H_{25}(C_6H_{10}O_4)O(C_6H_{11}O_5)$			
C12SUCR	*3.469*	$C_{11}H_{23}C(O)O(C_6H_{10}O_4)O(C_6H_{11}O_5)$			
C18SUCR	*5.292*	$C_8H_{17}CH=CHC_7H_{14}C(O)O(C_6H_{10}O_4)O(C_6H_{11}O_5)$			
Ethoxylated amines and amides (12 compounds)					
C11CONEO	*3.585*	$C_{11}H_{23}C(O)N(C_2H_4OH)_2$			
CnCONEmE		$C_nH_{2n+1}C(O)N[(C_2H_4O)NCH_3]_2$			
C9CONE3E	*2.299*	C11CONE2E	*3.398*	C11CONE4E	*3.611*
C9CONE4E	*2.193*	C11CONE3E **(v)**	*3.292*		
C12ALAE4	*3.413*	$C_{12}H_{23}NHCHCH_3C(O)O(C_2H_4O)_4H$			
C12GLYE4	*3.474*	$C_{12}H_{23}NHCH_2C(O)O(C_2H_4O)_4H$			
C12SARE4	*3.533*	$C_{12}H_{23}NCH_3CH_2C(O)O(C_2H_4O)_4H$			
C12AMEn		$C_{12}H_{25}CON(CH_3)CH_2CH_2O(C_2H_4O)_nH$			
C12AME3	*3.292*	C12AME6 **(v)**	*3.187*	C12AME9	*3.125*
Fluorinated linear ethoxylates and amides (20 compounds)					
CFnCONEm		$CnF_{2n+1}CH_2C(O)N[(C_2H_4O)_mCH_3]_2$			
CF6CONE3	*3.260*	CF8CONE3 **(v)**	*4.921*	CF10CONE	*6.523*
CF6SEn		$C^6F_{13}C_2H_4SC_2H_4O(C_2H_4O)_nH$			
CF6SE2	*4.602*	CF6SE5	*4.432*	CF6SE7	*4.319*
CF6SE3	*4.553*				
CF6SEnSEm		$C_6F_{13}C_2H_4(SC_2H_4OC_2H_4)_n(SC_2H_4OC_2H_4)_mOH$			
CF6SESE2	*4.638*	CF6SE2SE **(v)**	*4.585*	CF6SE3SE	*4.469*
HnE3		$H(CF_2)_nCH_2(C_2H_4O_3)CH_3$			
H4E3	*2.097*	H6E3	*3.523*		
FnE3		$F(CF_2)_nCH_2(C_2H_4O_3)CH_3$			
F4E3	*2.699*	F6E3 **(v)**	*4.097*		
FnC3NCOEm		$F(CF2)nCH_2CH_2CHNHCO(OC_2H_4)_mCH_3$			
F4C3NCOE2	*2.009*	F6C3NCOE2	*3.824*	F8C3NCOE2	*4.620*
F4C3NCOE2	*2.854*	F6C3NCOE3 **(v)**	*4.046*	F8C3NCOE3	*4.959*
Polyglycerol esters (11 compounds)					
GLYnOL-1		$C_{17}H_{33}COO(CH_2CHOHCH_2O)_mH$			
GLY4OL-1	*4.484*	GLY6-OL-1 **(v)**	*4.562*	GLY10OL-1	*4.676*
GLYnLA-1		$C_{11}H_{21}COO(CH_2CHOHCH_2O)_mH$			
GLY4LA-1	*4.402*	GLY6-LA-1 **(v)**	*4.446*	GLY10LA-1	*4.549*
GLYnST-1		$C_{17}H_{35}COO(CH_2CHOHCH_2O)_mH$			
GLY4ST-1	*4.650*	GLY6ST-1	*4.553*		
SORB-LA-1	*4.440*	$C_6O_4H_{11}OCOC_{12}H_{23}$			
SORB-OL-1	*4.578*	$C_6O_4H_{11}OCOC_{17}H_{33}$			
SORB-OL-3	*4.944*	$C_6O_4H_9(OCOC_{17}H_{33})_3$			
Carbohydrate-derivate esters, ethers and thiols (43 compounds)					
Cn-LACTOSE		$C_nH_{2n+1}COOC_{12}H_{23}O_{10}$			
C8-LACTOSE	*2.580*	C12-LACTOSE	*3.370*	C16-LACTOSE	*5.020*

Cn-LACTITOL		$C_nH_{2n+1}COOC_{12}H_{21}O_{10}$			
C8-LACTITOL	*2.561*	C12-LACTITOL **(v)**	*3.370*	C16-LACTITOL	*5.120*
N-C12-MPYR **(v)**	*3.740*	$C_{12}H_{23}OC_{12}O_{10}H_{21}$			
Cn-OCO-XYL		$C_nH_{2n+1}OCOC_5H_{11}O_4$			
C4-OCO-XYL	*0.921*	C6-OCO-XYL	*2.000*	C8-OCO-XYL	*2.357*
C5-OCO-XYL	*1.,237*	C7-OCO-XYL **(v)**	*1.745*	C9-OCO-XYL	*2.745*
Cn-O-XYL		$C_nH_{2n+1}OC_5H_{11}O_4$			
C4-O-XYL	*1.237*	C7-O-XYL	*2.036*	C10-O-XYL	*3.092*
C5-O-XYL	*1.420*	C8-O-XYL	*2.174*	C11-O-XYL	*3.523*
C6-O-XYL **(v)**	*2.027*	C9-O-XYL **(v)**	*2.678*		
Cn-S-XYL		$C_nH_{2n+1}SC_5H_{11}O_4$			
C4-S-XYL	*0.745*	C5-S-XYL **(v)**	*1.337*	C6-S-XYL	*1.796*
Cn-OCO-GLU		$C_nH_{2n+1}OCOC_6H_{11}O_5$			
C8-OCO-GLU	*2.796*	C16-OCO-GLU	*3.854*	C18-OCO-GLU	*3.699*
C12-OCO-GLU	*3.638*				
C12-O-MALT	*3.482*	$C_{12}H_{25}O(C_6H_{10}O_4)O(C_6H_{11}O_5)$			
C12CONE4	*3.301*	$C_{12}H_{25}CONH(C_2H_4O)_4H$			
C8TGLUPYR	*2.071*	$C_8H_{17}SC_6H_{11}O_5$			
BIS(CnGA)		$(C_nH_{2n+1}NC_3H_6NCOC_5H_{11}O_5)_2(CH_2)_2$			
BIS(C8GA)	*4.174*	BIS(C12GA) **(v)**	*5.420*		
BIS(C12GH)	*5.284*	$(C_{12}H_{25}NC_3H_6NCOC_6H_{13}O_6)_2(CH_2)_2$			
BIS(CnLA)		$(C_nH_{2n+1}NC_3H_6NCOC_{11}H_{21}O_{11})_2(CH_2)_2$			
BIS(C8LA)	*3.886*	BIS(C12LA)	*5.051*		
GLUPYR-n	(n=1,2)	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)]_2(CH_2)_{n+1}$			
GLUPYR-1	*2.143*	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)](CH_2)_2[CH_3(CH_2)_3OC_6O_5H_{10}CO)]$			
GLUPYR-2	*1.883*	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)](CH_2)_3[CH_3(CH_2)_3OC_6O_5H_{10}CO)]$			
GLUPYR-3	*2.669*	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)](CH_2)_2[CH_3(CH_2)_3OC_6O_5H_{10}CO)]$			
GLUPYR-4	*2.509*	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)](CH_2)_3[CH_3(CH_2)_3OC_6O_5H_{10}CO)]$			
GLUPYR-5	*1.801*	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)]C_6H_4[CH_3(CH_2)_3OC_6O_5H_{10}CO)]$			
GLUPYR-6	*0.959*	$CH_3(CH_2)_3OC_6O_5H_{11}$			
GLUPYR-7	*1.886*	$[CH_3(CH_2)_3OC_6O_5H_{10}CO)](CH_2)_2[CH_3(CH_2)_3OC_6O_5H_{10}CO)]$			

Table B.2 Marrero and Gani group definition and contributions after the regression
based on 150 experimental data of critical micelle concentration

1st Order Group (i)	C_i	2nd Order Group (j)	D_j	3rd Order Group (k) (new)	E_k
CH$_3$	-0.223	CH$_2$-CH$_m$=CH$_n$ *(m,n=0..2)*	-0.157	(CH$_2$)$_n$-(OCH$_2$CH$_2$)$_m$ *(n≤3)*	-0.456
CH$_2$	0.434	CHOH	-0.145	(CH$_2$)$_n$-(OCH$_2$CH$_2$)$_m$ *(n=5)*	-0.338
CH	1.009	CH$_m$(OH)CH$_n$(OH) *(m,n=0..2)*	0.012	(CH$_2$)$_n$-(OCH$_2$CH$_2$)$_m$ *(n>8, m≤1)*	0.581
CH=CH	0.000	CH$_{cyc}$-CH$_2$	0.524	(CH$_2$)$_8$-aCH	-0.879
aCH	0.273	CH$_{cyc}$-CH	0.000	(CH$_2$)$_n$-CH-OH *(n>9)*	0.472
aCH-CH$_2$	0.000	CH$_{cyc}$-OH	0.243	CH$_2$O-(CH$_2$CH$_2$0)$_n$-NHCO *(n=2)*	1.255
OH	-0.892	CH$_{cyc}$-O-	0.447	(CH$_2$)$_n$-CH$_2$CO *(n=9)*	-1.229
CH$_2$CO	0.324	CH$_{cyc}$-OOC	-1.267	(CH$_2$)$_n$-CH$_2$CO *(n=15)*	0.497
CH$_2$COO	-0.458	CH$_{cyc}$-CH$_2$	0.000	(CH$_2$)$_n$-CH$_2$CO *(n=2..4)*	-0.456
CH$_3$O	-1.060	AROMRINGs[1]s[4]	0.000	(CH$_2$)$_n$-CH$_2$CO *(n=6)*	2.370
CH$_2$O	-0.431			(CH$_2$)$_n$-CH$_2$CO *(n=14)*	0.393
CH-O	-0.048			(CH$_2$)$_n$-CH$_2$N *(n=6)*	0.462
aC-O	0.000			(CH$_2$)$_n$-OCH$_2$CHOH *(n<5)*	0.537
CH$_2$N	-1.732			(CH$_2$)$_n$-CH$_2$COO *(n=6..10)*	0.556
CONHCH$_2$	-1.116			(CH$_2$)$_n$-CH$_2$COO *(n=14..16)*	-1.921
CONCH$_3$CH$_2$	-1.225				
CON(CH$_2$)$_2$	0.438				
NHCO	-0.487				
CHF$_2$	-0.163				
CF$_2$	0.681				
CF$_3$	0.476				
OCH$_2$CH$_2$OH	-0.571				
OCH$_2$CHOH	-0.297				
CH$_2$S	-0.208				
CH$_2$ *(cyclic)*	0.000				
CH *(cyclic)*	0.303				
C *(cyclic)*	1.483				
O *(cyclic)*	0.000				
-O-	-1.219				

Table B.3 Prediction of the critical micelle concentration of tetraglycerol monostearate, after the introduction of new third order groups

Tetraglycerol monos-tearate	Molecular structure

Molecular formula:

$C_{30}H_{60}O_{10}$

First order groups	Occurrences	Group contribution
CH_3	1	-0.223
CH_2	19	0.434
OH	1	-0.892
CH_2CO	1	-0.324
OCH_2CHOH	4	-0.297

Second order groups	Occurrences	Group contribution
CHOH	3	-0.145
$CH_m(OH)CH_n(OH)$ (m,n in 0..2)	1	0.012

Third order groups	Occurrences	Group contribution
$(CH_2)_n$-CH_2CO (n=15)	1	-1.229

$$-\log(CMC)_{pred} = \sum_i N_i C_i + \sum_j M_j D_j + \sum_k O_k E_k = 4.609$$

$$-\log(CMC)_{exp} = 4.650$$

Appendix C: The virtual Product-Process Design Laboratory for Design and Analysis of Emulsified Formulated Products

In this appendix, the new template of the virtual Product-Process Design Laboratory for emulsified formulated product design is introduced. In order to carry out this specific design, the option "Emulsions" has to be selected from the main menu, as in Figure C.1.

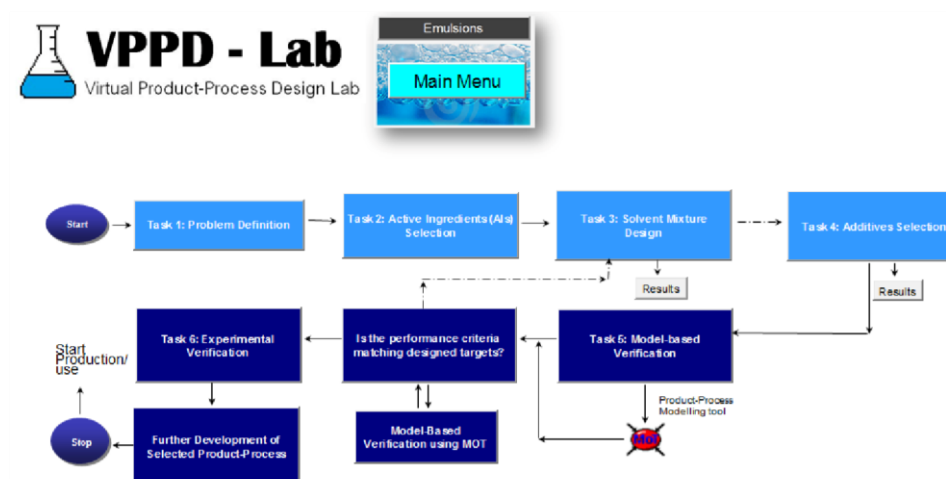**Figure C.1** Main menu of the VPPD-Lab

A template for emulsion design will be shown, as in Figure C.2.

Figure C.2 Template for the design of emulsified formulated products in the VPPD-Lab

165

Appendices

The case study relative to the hand-wash in the emulsified form (§7.3) is chosen to highlight the work-flow of the new template.

Task 1: Problem Definition

The option "Hand-wash" has to be selected, so that the relative knowledge base (§4.1) is retrieved. This way, the consumer needs are identified and translated into target properties: high foam-ability, non-irritability of the skin, cleaning performances, spread-ability, stability, safety and non-toxicity, as illustrated in Figure C.3.
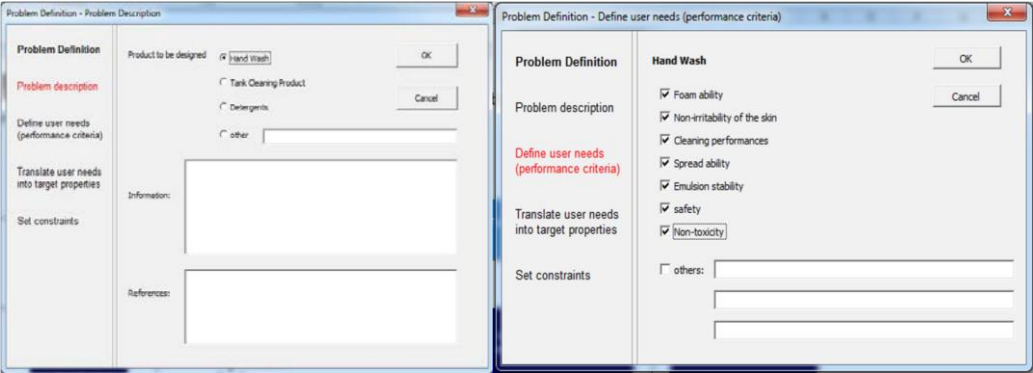

Figure C.3 User-interface for the problem definition in the VPPD-Lab, relative to the case-study of a hand-wash in the emulsified form

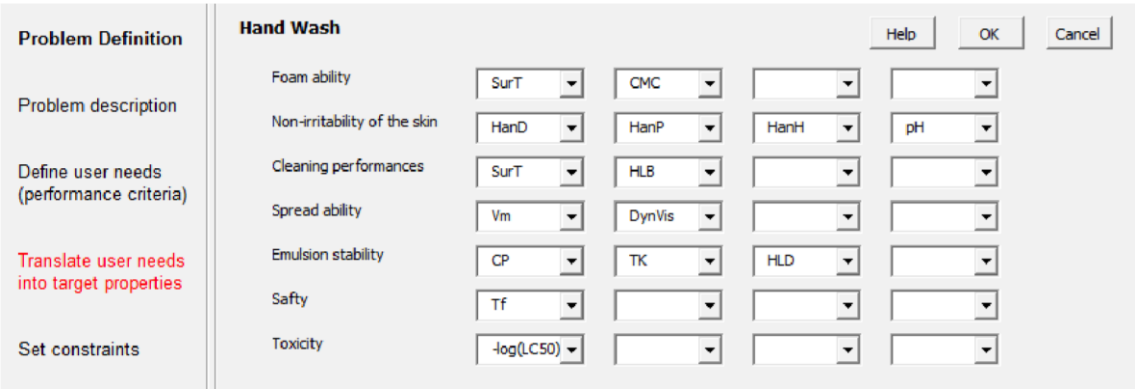The product needs are then translated into target properties, as in Figure C.4.

Figure C.4 User-interface for the definition of the target properties in the VPPD-Lab, relative to the case-study of a hand-wash in the emulsified form

Where SurT is the surface tension, HanD, HanP and HanH are the Hansen solubility parameters, Vm is the molar volume, and DynVis is the dynamic viscosity.
The numerical constraints are then generated from the knowledge base, but the software allows also manual input, as illustrated in Figure C.5.

Figure C.5 User-interface for the definition of the constraints on the target properties in the VPPD-Lab, relative to the case-study of a hand-wash in the emulsified form

Task 2: Active Ingredient Selection

The databases of ionic and non-ionic surfactants are retrieved, and the property constraints can be selected, together with the selection criteria, as in Figure C.6.
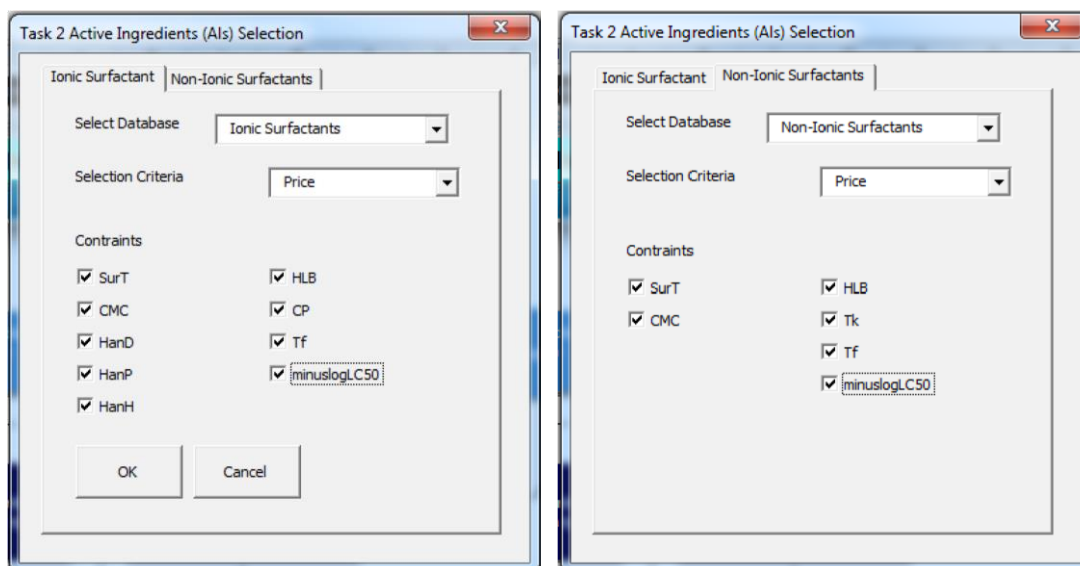
Appendices

Figure C.6 User-interface for the active ingredient selection in the VPPD-Lab, relative to the case-study of a hand-wash in the emulsified form

Task 3: Solvent Mixture Design

By clicking the button "Step 3", the software automatically set the boundaries on the mixture properties, select the necessary property models and optimize the composition of the emulsified solvent mixture. The results are summarized in Figure C.7.

Figure C.7 Results of the solvent mixture design in the VPPD-Lab, relative to the case-study of a hand-wash in the emulsified form

Task 4: Additives Selection

The user is asked to choose the phase in which the additives are desired to be soluble (aqueous or organic) and only the relative candidate ingredients are proposed by the software, as illustrated in Figure C.8.
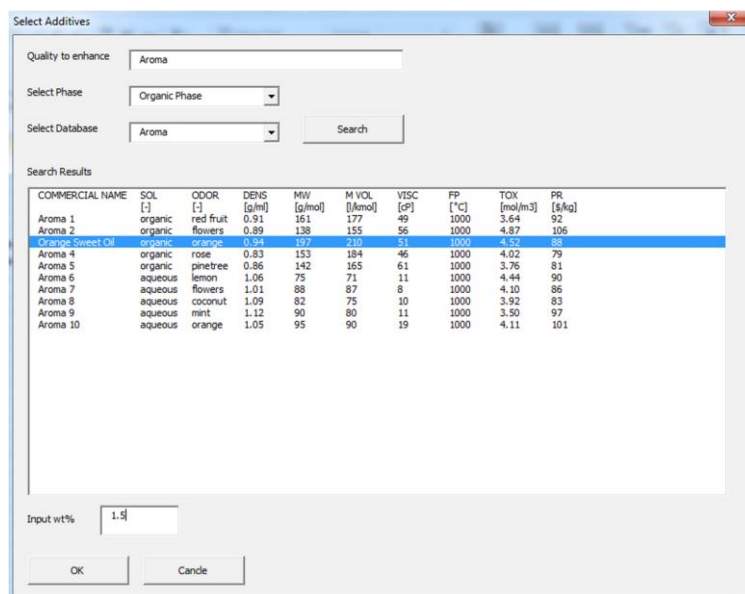
Figure C.8 User-interface for the additives selection in the VPPD-Lab,
relative to the case-study of a hand-wash in the emulsified form

As the final output of the template, the candidate composition of the emulsified formulated product is generated, and it is given in Figure C.9.

TASK 4 Additive Selection

	COMMERCIAL NAME	Conc.
		[we. %]
Non-ionic Surfactants	Tween 60	8.50
Ionic Surfactants	SDS	9.00
Water (Aqueous phase)	Water	54.19
Solvent (Organic phase)	Jojoba Oil	22.31
Aroma	Orange sweet aroma	1.5
Colorant	Orange colorant	1
Emollient	Polyquaternium 7	0.5
Preservative	Sodium benzoate	0.5
Co-surfactant	Propylene glycol	2.5

Figure C.9 Results of the additives selection in the VPPD-Lab, relative to the
case-study of a hand-wash in the emulsified form; the final formulation is proposed

Nomenclature

A	Parameter of the equations for the vapour pressure and for the heat of vaporization by Ceriani et al.
A	Descriptor of the equation for the Krafft temperature by Li et al.
A	Alcohol concentration
AAD	Average absolute deviation
A_i	Parameter of the UNIFAC model and of the equation for the surface tension by Suarez et al.
a_T	Parameter of the equation for the hydrophilic-lipophilic deviation by Salager
B	Parameter of the equations for the vapour pressure and for the heat of vaporization by Ceriani et al.
C	Cost
C	Parameter of the equations for the vapour pressure and for the heat of vaporization by Ceriani et al.
Ci	First-order group contribution
CMC	Critical micelle concentration
CP	Cloud point
c_T	Parameter of the equation for the hydrophilic-lipophilic deviation by Salager
D	Parameter of the equations for the vapour pressure and for the heat of vaporization by Ceriani et al.
D	Molecular dipole moment
D	Descriptor of the equation for the surface tension by Wang et al.
D_j	Second-order group contribution
DSD	Droplet size distribution
ECN	Effective carbon number
EHOMO	Descriptor of the equation for the critical micelle concentration by Wang et al.

171

Nomenclature

E_k	Third-order group contribution
ELUMO	Descriptor of the equation for the critical micelle concentration by Wang et al.
EON	Parameter of the equation for the hydrophilic-lipophilic deviation by Salager
ER	Evaporation rate
ET	Descriptor of the equation for the surface tension by Wang et al.
HLB	Hydrophilic-lipophilic balance
HLB	Hydrophilic-lipophilic deviation
IC	Descriptor of the equation for the Krafft temperature by Li et al.
KHO	Descriptor of the equation for the critical micelle concentration and for the surface tension by Wang et al.
KS2	Descriptor of the equation for the Krafft temperature by Li et al.
LC_{50}	Toxicity parameter
LD_{50}	Toxicity parameter
log(P)	Octanol-water partition coefficient
M	Viscosity ratio
M_j	Second-order group occurrence
MW	Molecular weight
N	Parameter of the equation for the dynamic viscosity by Sastri and Rao
n	Parameter of the UNIFAC mode and of the equation for the dynamic viscosity by Cao et al.
NBP	Normal boiling point
N_i	First-order group occurrence
nO	Descriptor of the equation for the Krafft temperature by Li et al. and for the surface tension by Wang et al.
O_k	Third-order group occurrence
p	Parameter of the UNIFAC mode and of the equation for the dynamic viscosity by Cao et al.
P	Pressure

P_{crit}	Critical pressure
$P98$	Descriptor of the equation for the Krafft temperature by Li et al.
PIT	Phase inversion temperature
P^{vap}	Vapor pressure
q	Parameter of the UNIFAC mode and of the equation for the dynamic viscosity by Cao et al.
R	Universal gas constant
S	Salinity
SD	Standard deviation
Ss	Descriptor of the equation for the open cup flash point by Gharagheizi et al.
T	Temperature
T^{90}	Evaporation time
T_{crit}	Critical temperature
T_f	Open cup flash point
T_k	Krafft temperature
T_m	Melting temperature
T_R	Reduced temperature
V_{crit}	Critical molar volume
$vEv1$	Descriptor of the equation for the open cup flash point by Gharagheizi et al.
V_m	Molar volume
X	Mole fraction
Z_{crit}	Critical compressibility factor
Z^{Ra}	Constant of the modified Rackett equation
A	Parameter of the equation for the hydrophilic-lipophilic deviation by Salager
Γ	Activity coefficient
δ_D	Dispersion Hansen solubility parameters
ΔG^{mix}	Gibbs energy change of mixing
δ_H	Hydrogen bonding Hansen solubility parameters

Nomenclature

ΔH^{form}	Heat of formation
ΔH^{fus}	Heat of fusion
ΔH^{vap}	Heat of vaporization
δ_P	Polar Hansen solubility parameters
δ_T	Hildebrand solubility parameters
ε	Dielectric constant
ζ	Generic thermo-physical property
θ	Parameter of the UNIFAC mode and of the equation for the dynamic viscosity by Cao et al.
μ	Dynamic viscosity
μ_B	Parameter of the equation for the dynamic viscosity by Sastri and Rao
v	Kinematic viscosity
ξ	Necessary class of ingredients
π	Parameter of the UNIFAC mode and of the equation for the dynamic viscosity by Cao et al.
ρ	Density
ρ_{crit}	Critical density
ρ_L	Liquid density
σ	Surface tension
Φ	Parameter of the UNIFAC mode and of the equation for the dynamic viscosity by Cao et al.
ψ	Volume fraction
ψ	Product attribute

References

Abildskov J. and O'Connell J.P. (2003). Predicting the solubilities of complex chemicals I. Solutes in different solvents. *Industrial and Engineering Chemistry Research*. 42, 5622-5634.

Abildskov J. and Kontogeorgis G.M. (2004). Chemical product design. A new challenge of applied thermodynamics. *Chemical Engineering research and Design*. 82, 1505-1510.

Abrams D.S. and Prausnitz J.M. (1975). Statistical thermodynamics of liquid mixtures: a new expression for the excess gibbs energy of partly or completely miscible systems. *AIChE Journal*. 21, 116-128.

Achenie L.E.L., Gani R. and Venkatasubramanian V. (2003). *Molecular design: theory and practice. Computer-Aided Chemical Engineering*, volume 12. Elsevier, Amsterdam, the Netherlands.

Affens W.A. and McLaren G.W. (1972). Flammability properties of hydrocarbon solutions in air. *Journal of Chemical and Engineering Data*. 17, 482-488.

Ansmann A., Gondek H. Kawa R. and Tesmann H. (2001). Sun screen agents. *US Patent 6280712*.

Arctander S. (1969). *Perfume and flavor chemicals. Aroma chemicals*. Allured, Carol Stream, IL.

Bagajewicz M. J. (2007). On the role of macroeconomics, planning and finances in product design. *AIChE Journal*. 53, 3155-3170.

Bagajewicz M., Hill S., Robben A., Lopez H., Sanders M., Sposato E., Baade C., Manora S. Hey Coradin J. (2011). Product design in price-competitive markets: a case study of a skin moisturizig lotion. *AIChE Journal*. 57, 160-177.

Bartok W. and Mason S.G. (1958). Particle motions in sheared suspensions. VII. Internal circulation in fluid droplets. *Journal of Colloid Science*. 13, 293-307.

Bernardo F.P. and Saraiva P.M. (2012). Integrated process and product design optimization: a cosmetic emulsion application. *Computer Aided Chemical Engineering*. 20, 1507-1512.

Cao W., Knudsen K., Fredenslund A. and Rasmussen P. (1993). Simultaneous correlation of viscosity and vapour-liquid equilibrium data. *Industrial Engineering Chemistry Research*. 32, 2077-2087.

Ceriani R., Gani R. and Meirelles A.J.A. (2009). Prediction of heat capacities and heats of vaporization of organic liquids by group contribution methods. *Fluid Phase Equilibria*. 283, 49-55.

References

Charpentier J.C. and McKenna T:F. (2004). Managing complex systems: some trends for the future of chemical and process engineering. *Chemical Engineering Science*. 59, 1617-1640.

Charpentier J.C. (2009). Perspective on multiscale methodology for product design and engineering. *Computers and Chemical Engineering*. 33, 936-946.

Chen C.C. (1996). Molecular thermodynamic model for Gibbs energy of mixing of nonionic surfactant solutions. *AIChE Journal*. 42, 3231-3245.

Cheng Y.S., Lam K.W., Ng K.M., Ko R.K.M. and Wibowo C. (2009). An integrative approach to product development – A skin care cream. *Computers and Chemical Engineering*. 33, 1097-1113.

Constantinou L. and Gani R. (1994). A new grou-contribution method for the estimation of properties of pure compounds. *AIChE Journal*. 40, 1697-1709.

Conte E., Martinho A., Matos H.E., and Gani R. (2008). Combined group-contribution and atom connectivity index-based methods for estimation of surface tension and viscosity. *Industrial and Engineering Chemistry Research*. 47, 7940-7954.

Conte E. Morales-Rodriguez R. and Gani R. (2009). The virtual product-process design laboratory as a tool for product development. *Computer Aided Chemical Engineering*. 26, 249-254.

Conte E., Gani R. and Malik T.I. (2010). The virtual product-process design laboratory to manage the complexity in the verification of formulated products. *Fluid Phase Equilibria*. 302, 294-304.

Conte E., Gani R. and Ng K.M. (2011). Design of formulated products: a systematic methodology. *AIChE Journal*. 57, 2431-2449.

Conte E., Gani R., Cheng Y.S and Ng K.M. (2011). Design of formulated products: experimental component. *AIChE Journal*. 58, 173-189.

Cussler E.L. and Moggridge G.D. (2011). *Chemical product design*. Cambridge University Press, Cambridge, ME.

Davies J.T. (1957). A quantitative kinetic theory of emulsion type, I. Physical chemistry of the emulsifying agent. *Proceedings of the Gas/Liquid and Liquid/Liquid Interface Conference*, 426-438.

Delgado E.J. and Diaz G.A. (2006). A molecular structure based model for predicting surface tension of organic compounds. *SAR and QSAR in Environmental Research*. 17, 483-496.

Duncan S. (1983). *Introduction to colloid and surface chemistry*. Butterworths.

Eden M.R. Jorgensen S.B., Gani R. and El-Halwagu M.M. (2004). A novel framework for simultaneous separation process and product design. *Chemical Engineering and Processing*. 43, 595-608.

Elbro H.S., Fredenslund A. and Rasmussen P. (1991). Group contribution method for the prediction of liquid densities as a function of temperature for solvents, oligomers and polymers. *Industrial and Engineering Chemistry Research*. 30, 2576-2582.

Fredenslund A., Jones R.L. and Prausnitz A.M. (1975). Group-contribution estimation of activity coefficient in non-ideal mixtures. *AIChE Journal*. 21, 1086-1099.

Gani R. and Pistikopoulos E.N. (2002). Property modelling and simulation for product and process design. *Fluid Phase Equilibria*. 194, 43-59.

Gani R. (2004). Chemical product design: challenges and opportunities. *Computers and Chemical Engineering*. 28, 2441-2457.

Gani R. (2005). Computer-aided methods and tools for chemical product design. *Chemical Engineering Research and Design*. 82, 1494-1504.

Gani R., Harper P.M. and Hostrup M. (2005). Automatic creation of missing groups through connectivity index for pure-component property prediction. *Industrial and Engineering Chemistry research*. 44, 7262-7269.

Gani R., Jimenez-Gonzalez C. and Constable D.J.C. (2005). Methods for selection of solvents for promotion of organic reactions. *Computers and Chemical Engineering*. 29, 1661-1676.

Gani R. and Ng K.M. (2014). Product design – From molecules to formulations to devices. *Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design*. 108-123.

Ghasemi J. and Ahmadi S. (2007). Combination of genetic algorithm and partial least squares for cloud point prediction of nonionic surfactants from molecular structures. *Annali di Chimica*. 97, 69-83.

Griffin W.C. (1949). Classification of surface-active agents by HLB. *Journal of the Society of Cosmetics Chemicals*. 1, 311-326.

Gross J. and Sadowski G. (2001). Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules. *Industrial and Engineering Chemistry Research*. 40, 1244-1260.

Gu T. and Sjoblom J. (1992). Surfactant structure and its relation to the Krafft point, cloud point and micellization: some empirical relationships. *Colloids and Surfaces*. 64, 39-46.

Guo C., Ren Y., Zhou J., Shao X., Yang X. and Shang Z. (2012). Toward a quantitative model and prediction of the cloud point of normal nonionic surfactants and novel Gemini

References

surfactants with heuristic method and Gaussian process. *Journal of Dispersion Science and Technology*. 33, 1401-1410.

Hancock B.C., York P. and Rowe R.C. (1997). The use of solubility parameters in pharmaceutical dosage form design. *International Journal of Pharmaceutics*. 148, 1-21.

Hansen C.M. (1967). *Hansen solubility parameters: a user's handbook*. CRC Press, Boca Raton, FL.

Harper P.M. (2000). *A multi-phase, multi-level framework for computer aided molecular design*. PhD dissertation, Technical University of Denmark, Denmark.

Harper P.M. and Gani R. (2000). A multi-step and multi-level approach for computer aided molecular design. *Computers and Chemical Engineering*. 24, 677-683.

Horvath A.L. (1992). *Molecular design*. Elsevier, Amsterdam, the Netherlands.

Hu J., Zhang X. and Wang Z. (2010). A review of progress in QSPR studies for surfactants. *International Journal of Molecular Science*. 11, 1020-1047.

Hukkerikar A.S., Sarup B., Ten Kate A., Abildskov J., Sin G and Gani R. (2012). Group contribution plus based estimation of properties of pure components: improved property estimation and uncertainty analysis. *Fluid Phase Equilibria*. 321, 25-43.

Hukkerikar A.S., Meier R.J., Sin G. and Gani R. (2013). A method to estimate the enthalpy of formation of organic compounds with chemical accuracy. *Fluid Phase Equilibria*. 348, 23-32.

Imhels E.C. and Gmehling J. (2003). Extension and revision of the group contribution method for the prediction of pure compound liquid densities. *Industrial and Engineering Chemistry Research*. 42, 408-412.

Israelachvili J.N., Mitchell D.J. and Ninham B.W. (1976). Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society, faraday transactions 2: Molecular and Chemical Physics*. 72, 1525-1568.

Joglekar G., Giridhar A. and Reklaitis G.V. (2014). A workflow modelling system for capturing data provenance. *Computer and Chemical Engineering*. In press.

Kahlweit M. and Strey R. (1985). Phase behaviour of ternary systems of the type water-oil-nonionic amphiphile (microemulsions). *Angewandte Chemie International Edition*. 24, 654-668.

Karayannis N.C., Mavrantzas V.G. and Theodorou D.N. (2004). Detailed atomistic simulation of the segmented dynamics and barrier properties of amorphous poly(ethylene terephthalate) and poly(ethylene isophtalate). *Macromolecules*. 37, 2978-2995.

Karunanithi A.T., Achenie L. and Gani R. (2005). A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures. *Industrial and Engineering Chemistry research*. 44, 4785-4797.

Katritzky A.R., pacureanu L.M., Slavov S.H., Dobchev D.A. and Karelson M. (2008). QSPR study of critical micelle concentration of nonionic surfactants. *Industrial and Engineering Chemistry Research*. 47, 9687-9695.

Klamt, A. (1995). Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *Journal of Physical Chemistry*. 99, 2224-2235.

Klein J.A., Wu D.T. and Gani R. (1992). Computer aided mixture design with specified property constraints. *Computers and Chemical Engineering*. 16, 229-236.

Klevens H.B. (1953). Structure and aggregation in dilute solutions of surface active agents. *Journal of American Oil and Chemistry Society*. 30, 74-79.

Kontogeorgis G.M., Voutsas E.C., Yakoumis I.V. and Tassios D.P. (1996). An equation of state for associating fluids. *Industrial and Engineering Chemistry Research*. 35, 4310-4318.

Kontogeorgis G.M. and Gani R. (2004). *Computer Aided Property Estimation for Process and Product Design*. Elsevier, Amsterdam, the Netherlands.

Landgren M. (1990). *Thermodynamic Modelling of Ionic Surfactant Systems*. PhD dissertation, Lund University, Sweden.

LeBard D.N., Levine B.G., Mertemann P., Barr S.A., jusufi A., Sanders S., Klein M.L., Panagiotopoulos A.Z. (2012). Self-assembly of coars-grained ionic surfactants accelerated by graphics processing units. *Soft Matter*. 8, 2385-2397.

Li J., Mundhwa M., Tontiwachwuthikul P. and Henni A. (2007). Volumetric properties, viscosities, and refractive indices for aqueous 2-(methylamino)ethanol solutions. *Journal of Chemical Engineering Data*. 52, 560-565.

Li X.S., Lu J.F. and Li Y.G. (2000). Study of ionic surfactant solutions by SAFT equation incorporated with MSA. *Fluid Phase Equilibria*. 168, 107-123.

Lin B.J. and Chen L.J. (2004). Liquid-liquid equilibria for the ternary system water+tetradecane+2-butoxyethanol. *Fluid Phase Equilibria*. 216, 13-20.

Marrero J. and Gani R. (2001). Group contribution based estimation of pure component properties. *Fluid Phase Equilibria*. 183, 183-208.

Martin M. and Martinez A. (2013). A methodology for simultaneous process and product design in the formulated consumer products industry: The case study of the detergent business. Chemical Engineering Research and Design. 91, 795-809.

Mattei M., Kontogeorgis G.M. and Gani R. (2012). A systematic methodology for design of emulsion based chemical products. *Computer Aided Chemical Engineering*. 31, 220-224.

References

Mattei M., Hill M., Kontogeorgis G.M. and Gani R. (2013). Design of an emulsion-based personal detergent through a model-based chemical product design methodology. Computer Aided Chemical Engineering. 32, 817-822.

Mattei M., Yunus N.A., Kalakul S., Kontogeorgis G.M., Woodley J.M. Gernaey K.V. and Gani R. (2014). The Virtual Product-Process Design Laboratory for Structured Chemical Product Design and Analysis. *Computer Aided Chemical Engineering*. 33, 61-66.

Mattei M., Kontogeorgis G.M. and Gani R. (2014). A comprehensive framework for surfactant selection and design for emulsion based chemical product design. *Fluid Phase Equilibria*. 362, 288-299.

Megnassan E., Le Goff D. and Proutiere A. (1994). Dielectric constant and density variations in pure liquids. Theoretical relations, comparison with experiments. *Journal of Molecular Liquids*. 59, 37-58.

Mitchell D.J., Tiddy G.J.T., Wairing L., Bostock T., McDonald M. (1983). Phase behaviour of polyethylene surfactants with water. *Journal of Chemical Society – Faraday Transactions*. 179, 975-1000.

Modarresi H., Conte E., Abildskov J., Gani R. and Crafts P. (2008). Model-based calculation of solid solubility for solvent selection – A review. *Industrial and Engineering Chemistry Research*. 47, 5234-5242.

Moggridge G.D. and Cussler E.L. (2000). An introduction to chemical product design. *Transaction IChemE-Part A*. 78, 5-11.

Ng K.M., Gani R. and Dam-Johansen K. (2007). *Chemical product design. Toward a perspective through case studies*. Elsevier, Amsterdam, the Netherlands.

Nielsen T.L., Abildskov J., Harper P.M., Papaeconomou I. and Gani R. (2001). The CAPEC Database. *Journal of Chemical Engineering Data*. 46, 1041-1044.

Peng D.Y. and Robinson D.B. (1976). New two-constant equation of state. Industrial and Engineering Chemistry Fundamentals. 15, 59-64.

Pilemand C. (2002). *Surfactants – Their abilities and important physico-chemical properties*. Report from Center for Miljø og Luftveje. Technical University of Denmark. Denmark.

Porter M.R. (1994). *The handbook of surfactants*. Chapman & Hall.

Prausnitz J.M. (1968). Thermodynamics of fluid-phase equilibria at high pressures. *Advances in Chemical Engineering*. 7, 139-206.

Rackett H.G. (1970). Equation of state for saturated liquids. *Journal of Chemical and Engineering Data*. 15, 514-517.

Ray A. (1971). Solvophobic interactions and micelle formation in structure forming nonaqueous solvents. *Nature*. 231, 313-334.

Reklaitis G.V. (2014). Perspective on systems engineering advances in process and product design. *Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design*. 1-4.

Ren Y., Zhao B., Chang Q. and Yao X. (2011). QSPR modelling of nonionic surfactant cloud points: an update. *Journal of Colloid and Interface Science*. 358, 202-207.

Renon H. and Prausnitz J.M. (1968). Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE Journal*. 14, 135-144.

Rosen M.J. (1984). *Surfactants and Interfacial Phenomena*. John Wiley and Sons. New York.

Salager J.L. (1996). Quantifying the concept of physico-chemical formulation in surfactant-oil-water systems – State of the art. *Progress in Colloid and Polymer Science*. 100, 137-142.

Sastri S.R.S. and Rao K.K. (1992). A new group contribution method for prediction viscosity of organic liquids. *Chemical Engineering Journal*. 50, 9-25.

Satyanarayana K.C., Abildskov J and Gani R. (2009). Computer-aided polymer design using group contribution plus property models. *Computers and Chemical Engineering*. 33, 1004-1013.

Sheldon T.J., Adijman C.S. and Cordiner J.L. (2005). Pure component properties from group contribution: Hydrogen-bond basicity, hydrogen-bond acidity, Hildebrand solubility parameter, macroscopic surface tension, dipole moment, refractive index and dielectric constant. *Fluid Phase Equilibria*. 231, 27-37.

Sjoblom J., Stenius P. and Danielsson I. (1987). Phase equilibria of nonionic surfactants and the formation of microemulsions. *Surfactant Science Series*, volume 23. Marcel Dekker, New York.

Slater J.C. (1951). A simplification of the Hartree-Fock method. *Physical Review*. 81, 385-390.Smith B.V. and Ierapepritou M. (2010). Integrative chemical product design strategies: reflecting industry trends and challenges. *Computers and Chemical Engineering*. 34, 857-865.

Soave G. (1972). Equilibrium constants from a modified Redlich-Kwong equation of state. *Chemical Engineering Science*. 27, 1197-1203.

Spencer C.F. and Danner R.P (1972). Improved equation for prediction of saturated liquid density. *Chemical Engineering Data*. 17, 236-241.

Suarez J.T., Torres-Marchal C. and Rasmussen P. (1989). Prediction of surface tension of nonelectrolyte solutions. *Chemical Engineering Science*. 44, 782-786.

References

Tsibanogiannis I.N., Kalospiros N.S. and Tassios D.P. (1994). Extension of the GCVOL method and application to some complex compounds. Industrial and Engineering Chemistry Research. 33, 1641-1643.

Van der Waals J.D. (1973). *Over de continuiteit van den gas- en vloeistoftoestand. PhD dissertation*. Leiden University, the Netherlands.

Van Speybroeck V., Gani R, and Meier R.J. (2010). The calculation of thermodynamic properties of molecules. *Chemical Society Reviews*. 39, 1764-1779.

Van Wesenbeck I., Driver J. and Ross J. (2008). Relationship between the evaporation rate and vapour pressure of moderately and highly volatile chemicals. *Bulletin of Environmental Contamination and Toxicology*. 80, 315-318.

Villadsen J. (1997). Putting structure into chemical engineering: proceeding of an industry/university conference. *Chemical Engineering Science*. 52, 2857-2864.

Voutsas E.C., Flores M.V., Spiliotis N., Bell G. Halling P.J. and Tassios D. (2001). Prediction of critical micelle concentration of nonionic surfactants in aqueous and non aqueous solvents with UNIFAC. *Industrial and Engineering Chemistry Research*. 40, 236-2366.

Wang Z.W., Li G.Z., Zhang X.Y., Wang R.K. and Lou A.J. (2002). A quantitative structure-property relationship study for the prediction of critical micelle concentration of anionic surfactants. *Colloid and Surfaces A: Physicochemical Engineering Aspects*. 197, 37-45.

Wilson G.M. (1964). Vapor-liquid equilibrium. XI. A new expression for the excess free energy of mixing. *Journal of the American Chemical Society*. 86, 127-130.

Wintermantel K. (1999). Process and product engineering – Achievements, present and future challenges. *Transaction of IChemE-Part A*. 77, 175-188.

Yamada T. and Gunn R.D. (1973). Saturated liquid molar volumes. The Rackett equation. *Journal of Chemical and Engineering Data*. 18, 234-236.

Yunus N.A., Gernaey K.V., Woodley J.M. and Gani R. (2014). Design of sustainable blended products using an integrated methodology. *Computer Aided Chemical Engineering*. 32, 835-840

Yunus N.A., Gernaey K.V., Woodley J.M. and Gani R. (2014). A systematic methodology for design of tailor-made blended products. *Computers and Chemical Engineering*. 66, 201-213.

www.capec.kt.dtu.dk

Computer-Aided Process-Product Engineering Center

Department of Chemical & Biochemical Engineering
Søltofts Plads
Building 227

DK-2800 Kgs. Lyngby
Denmark

Tel: (+45) 45 25 28 00
Fax: (+45) 45 88 22 58
E-mail: kt@kt.dtu.dk