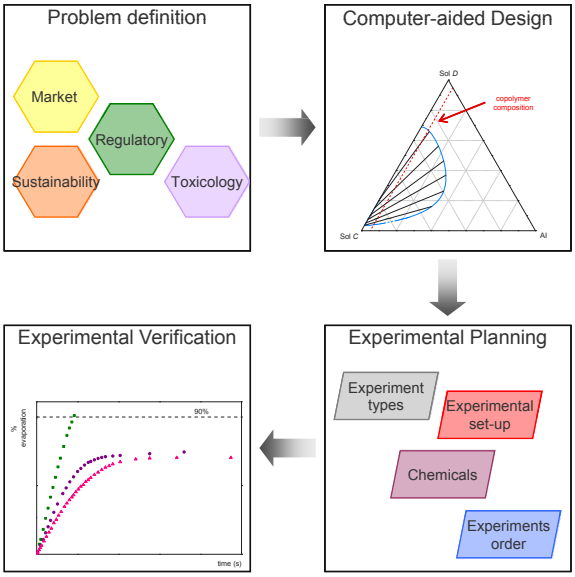
Innovation in Integrated Chemical Product-Process Design - Development through a Model-based Systems Approach

Conte, Elisa; Gani, Rafiqul; Abildskov, Jens

Publication date:
2010

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

Innovation in Integrated Chemical Product-Process Design -Development through a Model-based System Approach



Elisa Conte

Ph.D. Thesis

2010

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Innovation in Integrated Chemical Product-Process Design

Development through a Model-based System Approach

PhD Thesis

Elisa Conte

July 2010

Computer Aided Process and Product Engineering Center

Department of Chemical and Biochemical Engineering

Technical University of Denmark

ISBN-13: 978-87-92481-24-5
Printed by J&R Frydenberg A/S, Copenhagen, 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) at the Department of Chemical and Biochemical Engineering, from July 2007 to July 2010, under the supervision of Professor Rafiqul Gani. The project included an external research period at the Hong Kong University of Science and Technology (HKUST), Hong Kong. The project has been financed by a research grant from the Technical University of Denmark.

I would like to acknowledge my supervisor Rafiqul Gani for his collaboration, guidance and inspiring conversations throughout this project. I deeply believe that the opportunities I have been given during the three years PhD have profoundly affected my personality and broaden my perspectives.

I also owe special thanks to Professor Ka Ming Ng from the Hong Kong University of Science and Technology for hosting me in his group, providing new view points for my work and offering to use his laboratories and equipments. I am also grateful to Yuen S. Cheng (Alice) for her help in performing the experiments and in developing the experimental work-flow.

I wish to thank Akzo Nobel and specially Tahir I. Malik, for providing a challenging case study for my project and collaborating for publication.

Many thanks to Mark Juul, for the collaboration in the development of the extension of the software the 'virtual Product-Process Design laboratory' (virtual PPD-lab).

I would also like to acknowledge Grozdana Bogdanic, for her technical support with the FV-UNIQUAC model.

Special thanks to Marco Intelvi, for the development of the aroma database.

I am grateful to Professor Georgios M. Kontogeorgis for useful technical discussions.

To all CAPEC I also say thanks, for the smiling atmosphere I have been working in these three years.

"The scientific part of chemical engineering consists in breaking down complex systems which are then described using our understanding of fundamental phenomena…

… the engineering part consists in using the gained knowledge, even if incomplete, in the design of a product which matches the desired characteristics."

adapted from Wintermantel (1999)

Kgs. Lyngby, July 2010
Elisa Conte

Abstract

The ‘consumer oriented chemicals based products’ such as shampoos, sunscreens, insect repellents are used everyday by millions of people. They are structured products, constituted of numerous chemicals. This complexity gives the reason for which mainly experimental techniques are still employed in the design and verification of such products. The objective of this project is to tackle the problem with computer-aided tools at first, using experimental techniques for final testing, evaluation and amendment. In this way, time and resources can be spared and the product can reach the market faster and at a reduced cost.

The main contribution of this project is the development of an integrated methodology for the design and verification of formulated products. The methodology includes a first stage in which computer-aided techniques are employed to determine the base case product formula, a second stage in which experiments are planned and a third stage in which experiments are performed to validate the final product formula.

The main focus of the project is on the development of the computer-aided stage of the methodology described above. The methodology considers two different scenarios: the design of new products and the verification of modified and/or existing products. In the design scenario, since the identity of the chemicals belonging to the formulated product is unknown, and, thousands of design alternatives may be generated, the problem may encounter a combinatorial explosion unless appropriate model-based screening techniques are employed. In the verification scenario, a shortlist of candidate ingredients is provided, therefore the problem size is much smaller and rigorous property models can be employed/developed.

When using computer-aided tools for product design, several issues need to be addressed: new property models may need to be developed and/or the application range of existing property models may need to be extended (that is, new model parameters are needed), new and more efficient methods and tools for the application of the models may need to be developed, together with a flexible framework, which collects the methods and tools and allows their use in an integrated way. All these issues are addressed in this PhD project: new property models for the estimation of the target properties are developed; two algorithms for the design of binary mixtures and for the stability test of liquid systems are proposed, and the associated computer programs are also developed; the computer-aided stage of the methodology for formulation design and verification is implemented as an option in the software the ‘virtual Product-Process Design laboratory’.

Four case studies have been developed to illustrate the use of the proposed methodology. For two of these case studies the complete methodology has been applied, that is, including the stages of experimental planning and experimental testing/amendment. For the other two, only the computer-aided stage has been applied.

Resumé på dansk

Forbrugerorienterede kemiske produkter som shampoos, solcreme, insektsprays benyttes dagligt af millioner af mennesker. De er strukturerede produkter, ofte bestående af indtil flere forskellige kemikalier. Denne kompleksitet i sammensætning er årsagen til at eksperimentelle teknikker til stadighed hovedsagligt benyttes ved design og verificering af nye produkter. Formålet med dette projekt er at anvende computerbaserede værktøjer til først produktdesign og derefter eksperimentelle test som led i den endelige evaluering og forbedning af produktet. Derved kan resurser spares og det endelige produkt kan introducers på markedet hurtigere og billigere.

Hovedbidraget i dette projekt er at udvikle en integreret fremgangsmåde for design og verificering af produktformuleringer. Denne frengangsmåde inkluderer et indledende trin, hvori computerbaserede værktøjer benyttes til at opnå en mulig produktformel. I næste trin planlæges eksperimenter og i det tredje udføres disse for at finde den endelige produktformel.

Hovedvægten af projektet ligger i udviklingen af det computerbaserede trin i den integrerede fremgangsmåde. To scenarier for anvendelsen af fremgangsmåden behandles: Design af nye produkter eller verificering af forberinger for/eller af eksisterende produkter. I scenariet med design af nyt produkt skal tusinder af alternativer tages i betragtning, eftersom de enkelte ingredienser ikke er kendte. Det problem leder ofte til en kombinatorisk eksplosion, hvor brug af ”short cut” metoder er nødvendige. For verificeringsscenariet haves en begrænset liste af potentielle kandidat-ingredienser, hvorfor den kombinatoriske størrelse på problemet mindskes, og komplekse modeller for de fysisk-kemiske egenskaber kan blive anvendt/udviklet.

Ved anvendelsen af computerbaserede værktøjer til produktdesign skal flere forhold undersøges. Udvikling af nye modeller for de fysik-kemiske egenskaber kan være nødvendigt, eller anvendelsesområdet for eksisterende modeller kan være nødvendig at udvide, dvs. estimering af nye modelparametre. Nye og mere effektive metoder og værktøjer for anvendelse af disse modeller kan være nødvendige at udvikle sammen med en flexibel struktur, som tillader integreret anvendelse af metoderne og værktøjerne. Alle disse forhold er behandlet som led i dette arbejde: Nye modeller for fysisk-kemiske egenskaber er udviklet til at estimere de ønskede egenskaber, to algoritmer for design af binære blandinger og for beregning af fasestabilitet er udviklet med tilhørende programmel. Det computerbaserede trin i den overordnede fremgangsmåde er implementeret i en eksisterende ”in house” softwarepakke for produktdesign: ”virtual Product-Process Design laboratory”.

Fire specifikke studier har været udført som led i illustrering af den udviklede fremgangsmåde. For to af disse studier er alle tre trin i fremgangsmåden gennemført,

dvs. inklusive trinene med eksperimentplanlægning, evaluering og forbedring. For de resterende to studier er kun det første computerbaserede trin i fremgangsmåden blevet udført.

Contents

1

INTRODUCTION

In the last 50 years the horizon of chemical engineering narrowed to commodity chemicals. The research focused on perhaps 50 chemicals producing a powerful international industry based largely on petroleum feedstock (Cussler and Wei, 2003). In the present day the chemical industry is evolving beyond commodities towards specialty chemicals and 'consumer oriented chemicals based products' (Charpentier, 2010). It is not simply a change from commodities to specialties (Hill, 2004), but a substantial shift from material valued for their purity, to materials sold for their performance behaviour (Villadsen, 1997). The traditional oil and chemical companies are undergoing major changes and shifting their policy to high value added chemical manufacture (Westerberg and Subrahmanian, 2000). Costa *et al.* (2006) state that the chemical process industries have been facing dramatic social, economic and technical challenges, on a global and local scale. Consequently, they had to face rapid changes in the scope of their activities, in the strategies adopted to remain profitable and achieve sustainable growth.
Danckwerts (1966) foresaw this dramatic change in chemical engineering:

> *'It would be a great mistake to think of the content of chemical engineering science as permanently fixed. It is likely to alter greatly over the years, in response to the changing requirements of industry and to new scientific discoveries and ideas for their application.'*

The chemical product tree (shown in Fig. 1.1) gives an idea of the size of this shift from commodities to specialties. At the root of the tree there is a limited number of raw materials (~10,) which are processed to obtain the commodity products (~20). Specialty chemicals (~300) are produced from commodities. Finally, the leaves of the tree represent the 'consumer oriented chemical products'. They are obtained by processing and/or combining the chemicals of the previous product classes. The number of products grows exponentially from 10 for the raw material class, up to 30000 in the last class of 'consumer oriented chemicals based products'.

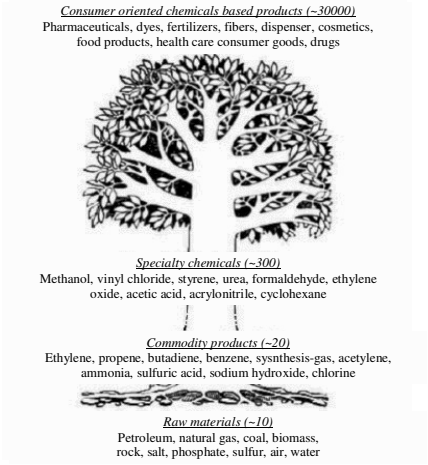


Figure 1.1. The chemical product tree. Classification of chemical-based products (adapted from: Gani, 2004a; Eden, 2003).

The 'consumer oriented chemicals based products' class is composed of formulations (or formulated products), devices and technology based consumer goods. To be commercially successful, these products need to satisfy the consumer needs as well as the environmental regulations, safety and health issues, which are becoming increasingly stricter. The 'consumer oriented chemicals based products' are needed because it is unlikely that a single molecule can satisfy all these needs, while mixing several chemicals together, many targets can be reached. The quality becomes a function of the properties: size, shape, colour, aesthetics, chemical and biological stability, degradability, therapeutic activity, solubility, mechanical, rheological, electrical, thermal, optical, magnetic characteristics for solids and solid particles, touch, handling, cohesion, friability, rugosity, taste, succulence, sensory properties, and so on (Charpentier and McKenna, 2004). Pharmaceuticals, paints, food, cosmetics, detergents, pesticides are included in this product class of complex and multicomponent systems, in which a presence of 5 to more than 20 ingredients is not unusual, and representing a range of different chemical compounds i. e. polymers, surfactants, solids, solvents, pigments, aromas and so on (Abildskov and Kontogeorgis, 2004).

Many of the chemical products of today and tomorrow ('consumer oriented chemicals based products') do not have much in common with those of twenty years ago in terms of molecular structure, number of ingredients involved, functions and structure. The chemical engineering science and practice must address this new reality, updating its skills and knowledge and evolving its research perspectives (Costa *et al.*, 2006).

Product design and engineering has recently been proposed as a possible third chemical engineering paradigm (Hill, 2009). A paradigm is a specific way of viewing

scientific reality, the mindset of a scientific community (Kuhn, 1996). The first paradigm in chemical engineering was stated in 1915, with the introduction of the unit operation concept (Favre *et al.*, 2002; Wei, 1996). With this mindset the engineers all over the world could recognize common elements in chemical processes. Chemical engineering became the study of unit operations giving the base for process design, which became a matter of deciding the sequence of unit operations. The second paradigm was introduced in the late 1950s (Committee on Chemical Engineering Frontiers, 1988; Favre *et al.*, 2002; Wei, 1996). It stated the importance of fundamental chemical and physical sciences (first principles) to overcome difficulties in solving important classes of problems. The second paradigm found its best exemplification in the book 'Transport phenomena' (Bird, Stewart and Lightfoot, 2002). Until recently, the chemical engineering community has ignored product design issues rather than purity, leaving to chemists the product development phase. The main objective of engineers had always been to minimize production costs.

Chemical product design is much more than a cost minimization. Product design aims at obtaining a product with added value through enhanced product qualities, and this is far more complex than profit maximization (Hill, 2009). Chemical product design and engineering has become a new mindset of chemical engineering.

Also Charpentier and McKenna (2004) recognize the importance of product design. In their opinion, the future of chemical engineering is supposed to face simultaneous research in four directions, where the synthesis of structured products combining several functions and properties required by the customer is one of them.

Coutinho *et al.* (2005) define chemical product design as a highly multidisciplinary area in which more emphasis is given on the product itself rather than the manufacturing process. Product engineering can also be thought as an alternative to process engineering in terms of finding innovative solutions.

Cussler and Moggridge (2001) define product design as the procedure by which customer needs are translated into commercial products. According to the perspective of Gani (2004a), chemical product design is that discipline that guides the developer in identifying the most appropriate chemical(s) that will exhibit and/or impart the desired behaviour.

Bagajewicz (2007) defines product design as a relatively virgin field, at least virgin from the use of tools and methods familiar to the PSE community. While process engineering has reached a high degree of scientific maturity, product engineering, is a younger, less mature area where the relationship between structure and properties still needs to be tackled, mainly at the molecular and microscopic scale (Wintermantel, 1999).

The development of methodologies, tools and strategies is crucial in order to organize, systematize and improve the design and evaluation of complex systems for the production of real products (Wintermantel, 1999). Common practice is still the

experiment-based trial-and-error approach. This design procedure is often inefficient with respect to time and resources and finding a feasible alternative may be due more to luck than to the ability of the designer (Gani, 2004b). Speeding up the product development is of paramount importance (Charpentier, 2009). Model-based computer-aided tools that perform a fast screening of numerous candidates and alternatives could provide several advantages to experimentation: the design procedure could be sped up; resources could spared since the number of candidates to verify experimentally would be drastically reduced; the optimal design could be reached since all possible candidates are screened. Gani (2004b) underlines the importance of collecting all methods and tools, as well as the gained knowledge, in a systematic framework through which the design and verification of chemical products is facilitated. Since chemical product design has a multidisciplinary nature, the systematic framework can be the "glue" that puts all the pieces of this discipline together.
Recently, various attempts have been made to develop systematic methodologies in the product design area. Computer-aided methods have been developed for:

- Molecular design (Harper and Gani, 2000);
- Solvent design (Gani and Brignole, 1983; Macchietto *et al.*, 1990; Pretel *et al.*, 1994; Gani *et al.*, 2005b);
- Mixture design (Eden *et al.*, 2004);
- Polymer design (Derringer and Markham, 1985; Vaidyanathan and El-Halwagi, 1994; Venkatasubramanian *et al.*, 1994; Satyanarayana *et al.*, 2009);
- Refrigerant design (Joback and Stephanopoulos, 1989; Churi and Achenie, 1996).

The algorithms employed are:

- 'Generate and test' algorithm (Gani and Brignole, 1983; Joback and Stephanopoulos, 1989; Pretel *et al.*, 1994; Constantinou *et al.*, 1996);
- Genetic algorithm (Venkatasubramanian *et al.*, 1994);
- Mathematical programming (Macchietto *et al.*, 1990; Vaidyanathan and El-Halwagi, 1994);
- Componentless design techniques (El-Halwagi *et al.*, 2000; Eljack *et al.*, 2007);
- Combinatorial optimization (Siddhaye *et al.*, 2004);
- Hybrid method (Harper and Gani, 2000).

The limitation of any computer-aided technique is closely related to the limitations of the property models being used, as well as to the complexity of the systems under investigation. But the words of Gani (2004a) have to be kept in mind:

> *'Although contribution alone from the PSE/CAPE community may not produce the magic chemical product, they will certainly help to find the magic solution, especially, in terms of getting the product faster and cheaper to the market.'*

The main objective of this thesis is to expand the CAPE capabilities to the design of complex products such as formulated products, which constitute a sub-group of the 'consumer oriented chemicals based products'. The problem is tackled keeping in mind the words of Wintermantel (1999), according to whom the 'scientific part' of chemical engineering consists in breaking down complex systems into subsystems, which are then described using our understanding of fundamental chemical and physical processes, while the 'engineering part' consists in using the gained knowledge (even if incomplete) in the design of a product that matches the desired characteristics, and a process that is capable of producing the desired product.

In this work, a systematic methodology for the design and verification of formulated products is proposed (in this work: the word 'verification' refers to the model-based verification that existing products/chemicals/chemical blends satisfy the required targets; the word 'validation' refers to experiment-based validation). The methodology consists of three stages: a computer-aided stage, a stage in which experiments are planned and a final stage where experiments are employed to test and amend the product, and thereby, reach the final product formulation. Each stage is divided into tasks, and every task is constituted of several sub-tasks, in which different methods and tools, models, databases and/or knowledge base are employed. The knowledge gained in each task is transferred to the successive one leading, finally, to the product recipe.

When the needed property models were not available in the open literature, they have been developed or adopted. Two methods and the corresponding tools have also been developed: one for the design of binary mixtures (MIXD) and the other for the stability test of liquid mixtures (STABILITY). An algorithm for the classification of mixtures is also proposed. Databases and knowledge base have been adapted/built. Case studies, illustrating the applications of the work-flow of actions, data and information between the sub-tasks, tasks, and stages, have been developed.

Another highlight of this project is the implementation of a systematic framework for the computer-aided design and verification of formulated products. The framework is based on the first stage of the systematic methodology (computer-aided stage), and collects the property models, methods and tools developed in this work in a systemathic and integrated manner. It provides a platform for performing virtual experiments on formulated products, or for the generation of a list of the most promising formulation candidates. The framework constitutes the extension of a in-house software, the 'virtual Product-Process Design laboratory' or virtual PPD-lab (Morales-Rodriguez, 2009).

1.1 Structure of the thesis

This PhD thesis is divided into eight chapters. The current chapter (Chapter 1) introduces the general product design problem and underlines the importance of focusing the attention in this relatively new area of chemical engineering. The objectives of this project are also given.

Chapter 2 concerns the theoretical background of product design, providing the reader with a detailed explanation of the objectives of this PhD work. Here, important concepts that constitute the pillars of the project are introduced: the general product design problem is defined; a classification of chemicals based products is given and the current solution approaches are considered; the issues and needs are highlighted, giving the basis for the different subjects discussed in this PhD work and that constitute the contents of the following chapters. In addition, the formulation design problem treated in this work is contextualized in the area of product design.

Chapter 3 highlights the modelling needs in formulation design. A formulated product is characterized by a certain performance or behaviour, which belong on the physicochemical properties of the pure compounds and/or mixture. All the physical and chemical properties needed in this work are discussed in this chapter, together with the issue of availability of the related property models. When property models are not available, they have been developed/adopted.

Chapter 4 is dedicated to the methods/tools specially developed in this work: an algorithm for the classification of mixtures according to hydrogen bonding properties, an algorithm for the design of binary mixtures (MIXD) based on the reverse approach and an algorithm for the stability test of liquid mixtures (STABILITY). Databases and knowledge base are also presented here.

In Chapter 5 the methodology for the design and verification of formulated products is presented. The problem decomposition in terms of tasks and sub-tasks is given, together with the flow of information. An introduction to the framework implemented in the in-house software, the 'virtual Product-Process Design laboratory', is also presented here.

Chapters 6 and 7 present the cases studies. All the case studies developed for the design scenario are gathered in Chapter 6: an insect repellent lotion, a sunscreen lotion, a paint formulation and, finally, a case study illustrating the use of the virtual PPD-lab. The case study developed for the verification scenario, a hair spray, is explained Chapter 7.

Chapter 8 is the conclusive chapter. Here a summary of the achievements is reported, together with the challenges and the future work in the field of formulation design.

2

PRODUCT DESIGN:

REVIEW & CHALLENGES

This chapter introduces general definitions and classifications commonly used in product design, which will often occur in the next chapters. In addition, this chapter contextualizes the PhD work in the vast area of product design:

- Which specific products are taken into consideration?
  - For these products, which solution approach is preferred?
    - Within this approach, which particular technique is employed?
- What is the contribution of this work with respect to the previous research?

At first, an appropriate definition for chemical product design is given (§2.1). Then a classification of products is proposed (§2.2), as well as a classification of solution approaches (§2.3). §2.4 is dedicated to the computer-aided design approaches. Computer-aided molecular design (CAMD) and computer-aided mixture/blend design (CAM^bD) are also evoked, in order to introduce the problem of formulation design. In §2.5, the issues and needs in product design are discussed. The contribution of this PhD work to the product design research area is presented in §2.6.

2.1 Problem definition

Moggridge and Cussler (2000) answer with a clear example to the question: what is chemical product design? They consider four chemicals based products: 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. There seems to be nothing in common for all these products. But, in fact, they are profoundly linked by the procedure by which they are designed. Moggridge and Cussler (2000) propose a four-level scheme for product design:

1. Level 1 (Needs). At first, the consumer needs have to be defined. Information about the consumer class that is going to use the product, the consumer desires as well as the regulatory, have to be obtained. All these requirements have then to be converted into quantitative specifications, which is an engineering task.

2. Level 2 (Ideas). Ideas to meet the needs have to be generated. In this step it is helpful to start from the industrial consensus that up to one hundred ideas are needed to get one successful product.
3. Level 3 (Selection). The best ideas previously generated are selected. Both qualitative matrix screening techniques and order-of-magnitude calculations are employed.
4. Level 4 (Manufacture). Finally, one has to decide what form the product should take and how it can be manufactured (process design).

According to Moggridge and Cussler (2000), chemical product design is this entire procedure. According to Ng, Gani and Dam-Johansen (2007) in chemical product design one tries to find a chemical product that exhibits certain functional properties. This practice involves the generation and screening of a large number of chemical molecules and/or mixtures of molecules, leading to a big combinatorial problem since the potential search space is very large. Chemical product-process design can be organized in 3 levels (Ng, Gani and Dam-Johansen, 2007):

1. Level 1 (Discovery). The desired product qualities are identified and then translated into chemical and physical properties (target properties). Based on this information, alternatives are generated, tested and evaluated in order to identify the product leading to the *a priori* defined characteristics.
2. Level 2 (Development). One of the products is selected and a process that can manufacture it is designed (process design).
3. Level 3 (Manufacturing and launch). Analysis, test and validation of the product and its corresponding process are performed.

In both the above definitions for product (-process) design common elements can be recognized: needs identification; translation of the needs into target properties; generation of ideas; selection of the best idea; process design; product manufacture; product launch. What is different is how these elements are grouped in the different levels.
It is worth noting that process design is considered as a part of product design in the definition of Cussler and Moggridge (2001), and that process design comes after product design in the definition of Ng, Gani and Dam-Johansen (2007). This is due to the fact that in process design the product to be produced has to be known, while in product design the product (and the corresponding manufacturing process) is not known.

2.1.1 Other useful definitions

Some of the terms to be used in the following chapters are defined below:

- *Formulated products*, or *formulations*. They constitute a class of 'consumer oriented chemical products'. They are formed by several ingredients (from 5 to 20). They can provide for several functions, and can have different forms (powder, solution, emulsion,…). For instance, a sunscreen lotion has the function of blocking the UV radiation, avoiding skin cancer, slowing the skin aging. Sunscreens can have the form of creams (emulsions) or solutions of oils, which can also be sprayed through a nozzle.
- *Performance criteria*, *consumer needs* or *product attributes*. They are also referred to as 'consumer preferences'. They are the requirements for a certain product when sold in the market. Consumers define most of these criteria. For instance, a cosmetic product that gives stickiness on the skin is not pleasant, therefore, it is unlikely be a successful product. Also the regulatory/law poses some constraints, such as the VOCs (Volatile Organic Compounds) emissions for a spray product.
- *Main product function,* or *activity.* It is the main function of the product, the reason for which consumer buy the product. For an insect repellent, for instance, the main function is to repel mosquitoes. A product can also have more than one main function, for instance a sunscreen has the functions of protecting the skin from sunburns and skin cancer, but also of preventing the skin aging.
- *Target properties,* or *quality factors.* When solving a product design problem, the consumer needs have to be translated into physicochemical properties, that is, the target properties. Each performance criteria can be represented by one or more properties. In the case of a paint formulation, consumers prefer a product that is easy to spread on the wall. The physicochemical properties (target properties) that affect spread-ability are, for instance, viscosity, surface tension and density.
- *Active Ingredient* (AI), or *key ingredient.* The AI is the ingredient that provides the main function of a formulation (also referred to as 'activity'). Since a formulation can provide several functions, more than one AI may be present in a single formulated product. In the case of a sunscreen lotion, for instance, UV blockers are needed to filter the sun radiation and avoid skin cancer, while antioxidants are needed to slow the aging of the skin.

2.2 Products classification

Based on the classification of chemical products given in Chapter 1 and the chemical product tree (Fig. 1.1), each product type is briefly explained below:

- Commodity chemicals, obtained from the processing of the raw materials (oil, gas,…) in very large quantities. They are sold on the basis of their purity.

- Specialty chemicals, pure compounds produced in smaller quantities than the commodities. They are sold on the basis of a specific benefit or function.
- 'Consumer oriented chemical products', which include formulated products (formulations), devices and technology based consumer goods:
  - Formulated products: a good example is cosmetics and food consumer goods. They can be defined as combined systems since they are constituted of several ingredients and they are designed to meet the consumer needs. They are often multifunctional since they can carry out more than one function and they can be micro-structured;
  - Devices: they carry out a physical or chemical transformation, such as a polymeric microcapsule for the controlled release of ingredients;
  - Technology based consumer goods: for instance, a transdermal patch or a disposable diaper. The functionality of these products is provided by a chemical/physical technology.

Costa *et al.* (2006) propose an unnecessary extension of the product classification, including also bio-based products and/or concepts, such as innovative biomaterials, drugs, tissues and metabolic engineering technologies, and virtual chemical products, such as softwares to simulate chemical processes (Aspen Plus®, ProII®).
As mentioned in Chapter 1, traditional chemical engineering has been focusing on commodity chemicals. These chemicals are characterized mainly by their purity. Process design has developed around this kind of chemicals, with the objective of reaching an efficient manufacture to minimize the production cost, hence, the product price. The process is usually continuous, optimized and energy integrated. The introduction of more complex products (formulated products, devices,…) has shifted the focus of chemical engineering to the product design field. For these products, process efficiency is less important, and they are usually produced in generic batch equipments. They are valued for their special functions, rather than for their efficient manufacture.
As previously stated, this PhD project concerns the 'consumer oriented chemical product' class, and in this class, formulated products are considered. Formulations are usually constituted of three different classes of compounds/mixtures:

- One or more key ingredients responsible for their functionality, which will be referred to as the *Active Ingredients* (AIs);
- Some supporting ingredients for enhancing the product performance, promoting some product qualities and so on, which will be referred to as *additives*. Additives are usually present in low concentrations;
- A delivery system, responsible of delivering the AI/AIs on the desired surface. Depending on the application, different delivery systems can be chosen to

deliver the AIs. These delivery systems can also have different forms, from solid composites to aerosols. Wibowo and Ng (2002) propose a classification of the various product forms and delivery systems. Table 2.1 shows this classification along with examples in three major application areas: cosmetics and personal care, health care and pharmaceuticals, households and office supplies.

Solid products can have their own shape such as composites, solid foams, tablets and capsules (Fung and Ng, 2003), or they can be powders (or granules). Semi-solid products can be pastes, if containing a large portion of solids, and creams, if containing immiscible liquid phases (Wibowo and Ng, 2001; Cheng *et al.*, 2009). Liquid products include single phase liquids (solutions), as well as dispersions of solid in liquid (suspensions), liquid in liquid (emulsions), and gas in liquid (foam). Macromolecule solutions are solutions of large molecules such as proteins, polymers, and surfactants. Formulations can also assume a gaseous form, the aerosols.

A suspension can also be obtained by suspending a solid in a solution. An aerosol can also be obtained adding a propellant to a liquid solution. It has to be underlined that solutions and diluted suspensions can also be sprayed, without the addition of a propellant (for instance, a perfume).

The types of formulations considered in this work are characterized by a liquid delivery system (liquid formulations), therefore they are solutions. Diluted suspensions and aerosols are also investigated, keeping in mind that these two kinds of formulations are an extension of the solution type of formulated products.

The performance of formulated products, and therefore, the customer satisfaction depends on product functions that are related to two factors: material properties and product microstructure. The material properties are physical and chemical properties, such as density, viscosity, surface tension and so on. The product microstructure describes how the formulation ingredients are assembled. It is related to intrinsic thermodynamics of the system, molecular weight distribution, phase volume fraction, polymorphism, and so on. While the material properties strictly depend on the choice of the ingredients, the product microstructure may also depend on the product processing and the operating conditions (Wibowo and Ng, 2002).

Table 2.1. Classification of formulated products and relative examples in three major application areas (Wibowo and Ng., 2002).

Physical form		Product form and delivery system	Examples		
			Cosmetic and personal care	Health care and pharmaceuticals	Households and office supplies
SOLID	SHAPED	Composites	Bar soap, lipstick	Inhalant stick	Compact disk, glue
		Capsules	-	Whale oil capsule	Microencapsulated carbon less capsules
		Tablets	-	Aspirin	Moth balls
		Solid foams	-	-	Styrofoam
	BULK	Powders and granules	Facial powders, baby powders	Powdered herbal medicine	Powdered detergent, dry toner
SEMI-SOLID		Pastes	Toothpaste	Pain relief ointment	Silicone sealant, metal adhesive
		Creams	Cleansing cream, hair cream	Pharmaceutical cream	Multipurpose adhesive
LIQUID		Liquid foams	Shaving foam	-	-
		Macromolecular solutions	Mouthwash, shampoo	-	Dishwashing liquid
		Dilute emulsions and suspensions	Suntan lotion, nail polish	Penicillin	Correction fluid, writing ink
		Solutions	Perfume	Eye drop, ginseng extract	Drain cleaning solution
GAS		Aerosols	Hair spray	Sore throat spray	Aerosol paint, antifreeze spray

2.3 Solution approaches

The solution approaches to product design can be classified in the following types (Ng, Gani and Dam-Johansen, 2007):

- Experiment-based trial-and-error approach: this approach is employed when mathematical models for the estimation of the target properties are not available. A large number of consumer products are developed through trial-and-error experiments. Since the desired properties need to be measured, not many candidate products are normally considered. Past knowledge and experience are crucial in this approach.
- 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. This approach is able to find feasible candidates within the application range of the models. Model-based computer-

aided molecular design (CAMD) and blend design (CAM^bD) 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. At the outer level, predictive models with wide application are employed. As one goes from the outer levels to the inner levels, the number of candidates decreases. The inner levels employ correlations and/or experiments.

Model-based and experiment-based trial-and-error approaches are compared in Table 2.2 in terms of tools employed, objectives, research environment and uncertainty. Main advantages and disadvantages are also highlighted.

Table 2.2. Comparison between model-based approach and experiment-based approach.

	model-based	**experiment-based**
Tools	mathematical predictive models	knowledge base, experience of few expertises
Objective	screen numerous candidates, obtain a base case	manufacture the end-use product
Environment	virtual reality	reality
Uncertainty	models uncertainty, assumptions, hypothesis	no uncertainty
Advantages	time and resources are spared	manufacture of end-use products
Disadvantages	high uncertainty, necessity of an experimental validation, some properties cannot be modelled (scent, appearance,…), models are limited to some kind of chemical	long development times, high consumption of resources, necessity of the knowledge base

The experiment-based trial-and-error approach has the objectives of characterizing the properties of the ingredients as well as the product prototypes, to verify if these properties match the requirements and to change the product composition until the requirements are satisfied. This approach necessitates the use of a knowledge base and/or industrial expertise in terms of heuristics and guidelines, in order to adjust the product attributes to a target value. Most of the times there is experimental evidence that, for instance, chemical A is better than chemical B for a specific role in the product. However, no explanation is given or no further investigation is performed to try to investigate the underlying phenomena, mainly for lack of time.
The model-based approach aims at screening numerous product alternatives in order to identify a small number of candidates that could be further tested, validated and amended through experimental research. The uncertainty of the computer-aided method depends on the reliability of property models, on the assumptions and on the hypothesis employed. In addition, modelling cannot cover a certain class of target properties that are crucial for the kind of product being considered here: the cosmetic and sensorial factors, such as the turbidity/colour of the product, the scent, the greasiness and stickiness (properties that can be easily measured through experiments). With the computer-aided approach every choice of chemical and every

result is moved by considerations based on an understanding of fundamental phenomena.
The experiment-based trial-and-error approach is very demanding in terms of time and resources (in fact it is difficult to predict the total time needed to reach the optimal product), while through the computer-aided approach time and resources in the experimental phase can be spared (the prototype is close to optimization level and small changes are expected to modify the performance).
The integrated approach combines modelling and experiments, therefore, the uncertainties of the model-based approach are compensated by the experimental part while the number of experiments is reduced through model-based predictions. Mathematical models are used to generate and test alternatives in order to identify a small number of candidates that will be further investigated through more rigorous models, collected data and/or experiments. Therefore the search space is reduced and time and resources can be spared. The expensive experimental validation is reserved only for the most promising candidates.
The methodology proposed in this PhD thesis is based on the integrated approach.

2.4 Computer-aided product design

A general product design problem is formulated as follows: given a set of desired specifications, for example, chemical/physical properties for a product, determine the chemical product that satisfies the *a priori* defined targets. If the product is a pure compound, the design problem is referred to as Computer Aided Molecular Design or CAMD. If the product is a mixture or a blend, the design problem becomes a Computer Aided Mixture/blend Design or CAM^{b}D. If the product is a formulated product, the problem is known as formulation design problem. Keeping in mind the above problem definition, chemical product design can be described as 'reverse property prediction problem' (Gani and Pistikopoulos 2002; Eden, 2003). Fig. 2.1 highlights the difference between a property prediction problem and a product design problem (Gani, 2004b).

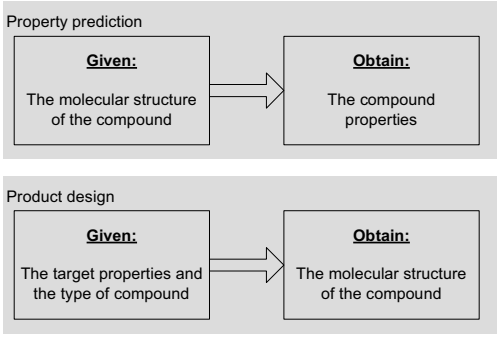


Figure 2.1. Comparison between property prediction and chemical product design (Gani, 2004b).

In a property prediction problem the chemical structure of the compound is known and the properties are calculated through property models. In a product design problem, this is reversed: the desired compound properties are known and the chemical structure of the compound needs to be identified.
In most of the solution methods employed in computer-aided chemical molecular design (CAMD) and mixture/blend design (CAM^bD), property models are employed in a 'generate and test' solution approach where the property prediction problem is solved repeatedly to test the generated alternatives.

2.4.1 Computer Aided Molecular Design-CAMD

Molecular design problems can be defined as follows: given a set of building blocks and a specified set of target properties, determine the molecule or molecular structure that matches these properties (Gani, 2004b). The molecular structure of a compound is usually represented through groups (Harper *et al.*, 1999) and/or connectivity indices (Camarda and Maranas, 1999). The methods employed follow the same main steps: generate feasible chemical structures, estimate the physicochemical properties through property models (or measurements), select the molecules that match the desired targets (reject the molecules that do not match).
The methods employed can be classified as:

- Generate and test method (Gani *et al.*, 1991; Joback and Stephanopoulos, 1989; Pretel *et al.*, 1994; Constantinou *et al.*, 1996; Friedler *et al.*, 1998; Harper and Gani, 2000): all the steps of the generate-predict-select procedure are performed sequentially. The generation step is performed using a combinatorial approach. The problem here is to avoid the so called combinatorial explosion, which occurs when the size of the problem becomes so large that computational time becomes too excessive;
- Mathematical programming: all the steps of the generate-predict-select procedure are performed simultaneously. The molecular design problem is formulated as an (optimization) problem where the constraints are treated as mathematical equalities and/or inequalities and the performance indices are combined into an objective function, which is minimized through an appropriate numerical method. These problems are usually solved by optimization methods, such as the Mixed Integer Non Linear Programming (MINLP) solution methods (Odele and Macchietto, 1993; Duvedi and Achenie, 1996; Camarda and Maranas, 1999). CAMD algorithms based on Mixed Integer Linear Programming (MILP) solution methods have also been proposed. Through this formulation is always possible to find the global optimum (product candidate), but it requires the use of linear models for the description of properties. This can be achieved by using linear models or linear approximation to a non-linear model (Maranas, 1996; Raman and Maranas, 1998);

- Stochastic optimization (Marcoulaki and Kokossis, 1998; Venkatasubramanian *et al.*, 1995): it is based on successive pseudo-random generation of solution alternatives;
- Decomposition methods (Solvason *et al.*, 2009; Chemmangattuvalappil *et al.*, 2010; Karunanithi *et al.*, 2005): the problem is decomposed in sub-problems and different tools are employed for each sub-problem.

Database search may also be employed. It involves the selection of known compounds (Joback and Stephanopoulos, 1994; Modi *et al.*, 1996) from a database. This method does not involve any design since no new molecules are generated, therefore is not listed in the above set of computer-aided methods for molecular design problems.

2.4.2 Computer Aided Mixture/blend Design-CAM^{b}D

Mixture/blend design problems can be defined as follows: given a set of chemicals and a set of property constraints, determine the optimal mixture and/or blend (Gani, 2004b). The chemicals to be mixed together are unknown, and also their relative compositions in the blend are not known. But the molecular structures of the candidate chemicals are known.
Mixture design is similar to molecular design in the sense that both design problems combine building blocks in order to reach some *a priori* defined targets: in molecular design the building blocks are the groups (CH_3, CH_2, OH,…) or atoms (C, H, O,…), while in mixture design, the building blocks are molecules. But mixture design is more challenging than molecular design for the following reasons:

- Mixture design requires also the calculation of the relative amounts of chemicals to blend together (concentration);
- Mixture design implies the need to handle phase behaviour issues, that is, miscibility/solubility between the ingredients.

According to Gani (2004a) mixture/blend design is still a quite immature area, and there is just limited knowledge and know-how about a systematic approach for the design and verification of this type of chemical products. The main efforts have been directed to the design of solvent mixtures (Sinha *et al.*, 2003; Klein *et al.*, 1992; Karunanithi *et al.* 2005).

2.4.2.1 Formulated products design

The design of formulated products can be considered a type of mixture/blend design (CAM^{b}D), where the number of compounds may be high (5-20), and the ingredients are quite different between each others, not just solvents. In chemical formulations, solvents, polymers, pigments, surfactants, aroma compounds, and so on are blended together. In addition, these products also have complex structures (polymers,

pharmaceutical ingredients,…) and forms (emulsions, suspensions,..). This factor further complicates the design.
Given the complexity of the formulation design problem, methods which can manage this complexity need to be employed. The method developed in this PhD project recall the decomposition method used in CAMD problems (Karunanithi *et al.*, 2005): the formulation design problem is decomposed in sub-problems, each sub-problem is solved separately using different techniques. For example, some sub-problems are solved through database search, while other sub-problems are solved through mathematical (linear) programming techniques.

2.5 Issues and needs

The research issues and needs currently faced related to CAM^{b}D are many and diverse. They can be organized under the following generic points:

- Problem definition (Harper, 2000; Gani, 2004a; Costa *et al.*, 2006);
- Property models (Gani, 2004a; Costa *et al.*, 2006);
- Methods and tools (Gani, 2004a; Costa *et al.*, 2006);
- Methodologies (Gani, 2004a ; Costa *et al.*, 2006);
- Systematic frameworks (Gani, 2004a; Costa *et al.*, 2006);
- Multidisciplinary modelling (Charpentier J. C., 2002; Charpentier and McKenna, 2004; Bagajevicz, 2007);
- Multiscale modelling (Charpentier J. C., 2002; Charpentier and McKenna, 2004; Gani 2004b; Morales-Rodriguez, 2009).

2.5.1 Problem definition

The reliability of a solution to a product design problem (CAMD, CAM^{b}D and formulation design) largely depends on the problem definition. This step consists of identifying the needs for a specific product, and relating these needs to physicochemical properties. There is the necessity of developing knowledge-based systems that may guide the chemical product designer to convert the problem representation space from customer needs to technical specifications, as well as to specify their target (goal) values for a large range of chemical product design (Harper, 2000; Gani, 2004a). Costa *et al.* (2006) claim this is to be relevant for improving the understanding of the relationship between product performance, product composition, ingredients properties, processing variables and usage variables.

2.5.2 Property models

Models play a central role in the solution of all computer-aided product design problems, since the reliability of the solution to such problems is strongly affected by

the choice of the model, the uncertainties in property estimations, the availability of model parameters and the size of the search space.
The most significant limitations to the use of property models are associated with the unavailability of model parameters and the accuracy of prediction. If model parameters are not available for a generated molecule and/or a physicochemical property, this molecule can no longer be considered as one of the candidate product alternative since its properties cannot be estimated (in CAM^{b}D, this molecule cannot be considered as a mixture component). This may eliminate a potentially optimal molecule. The major need here is to develop/adopt property estimation models with few parameters but with wider application ranges. That is, make them truly predictive and, at least, qualitatively correct.
In CAM^{b}D and formulation design solubility and miscibility issues are very important target properties. When the problem involves small and simple molecules such as solvents, group-contribution methods are usually sufficient to calculate physicochemical properties and to predict the solubility. When the molecules are complex and large such as multi-functional molecules or polymers, higher level property modelling is needed. Models for prediction of properties of structured formulations, colloidal dispersions, emulsions, chiral separations, etc., are not currently available in a form that can be implemented as part of a CAM^{b}D method. In addition, these models have been used for specific applications and the models parameters available are restricted only to few systems.

2.5.3 Method and tools

Design algorithms that do not focus primarily on the product cost but that take into consideration the various aspects and implications of product design need to be developed. The objective function should be formulated to address product and process performance and consider not only economic issues, but also risk analysis, uncertainty, environmental impacts (i. e., VOCs emissions), quality costs and health (i. e., toxicity), safety (i. e., flammability) and social concerns (customers satisfactions) over the entire chemical product life cycle. Flexible solution strategies are also necessary. A new class of computer-aided methods and tools that is systematic but flexible, simple but accurate is needed. In addition, it should be possible to create the necessary models for a given problem. It has been previously underlined that product design is a reverse problem compared to property prediction. Flexible solution approaches should also be able to solve these problems with the reverse approach (Eden *et al.*, 2004; Gani and Pistikopoulos, 2002).

2.5.4 Methodologies

Products have traditionally been developed through costly and time consuming trial-and-error design procedures. The development of systematic methodologies, with the related work-flows and data-flows for the inter-related activities involved in the

design of new products has been recognized as one of the main research challenges in the context of chemical product engineering.

2.5.5 Frameworks

The solution of a wide range of chemical engineering problems requires a suite of different methods and tools. For instance, an application involving the design/verification of a formulated product requires database search, property prediction, determination of phase diagrams, sensitivity analysis, mixture/blend design, simulation of product performance and many more steps. Developing methods and tools for product design problems is almost as important as being able to use them in a integrated manner, allowing inter-changes of information, data and results. The architecture of the framework should be flexible, allowing the addition of new models, data and adaptation of existing models for future extension of the software application range. The software should also be able to capture past experience through the further extension of databases and/or the creation of knowledge based libraries. In addition, a user-friendly interface is required in order for the software to have potential industrial application, and to provide also a significant contribution to the effective teaching of chemical product engineering.

2.5.6 Multidisciplinary modelling

Pricing and microeconomics, as well as supply chain, process synthesis and finances are needed when one wants to design new products. The systematic frameworks for product design should take into account not only product composition, structure and functionality, but also the manufacturing investment and costs, the associated supply chain and the consumer behaviour with respect to the product price (Bagajewicz, 2007). Multidisciplinary approaches need to be developed in response to the increasing environmental, societal and economic requirements and to the transition towards sustainability, that is, environmental protection, security, societal demands, and business including better conversion and selectivity of raw materials and energy for consumer desired product quality (Charpentier and McKenna 2004). Considering the multidisciplinarity in product design would return the discipline closer to practices in industry.

2.5.7 Multiscale modelling

It is necessary to organize time/length scales and complexity levels in product and process engineering in order to: first, understand and describe the phenomena and the properties at nano-, micro- and meso-scales; second, to understand the relationships between the different scales. The overall objective is to use the gained knowledge to design better molecules/blends and processes.

2.6 Addressing issues and needs

This PhD project will address some of the issues and needs highlighted in §2.5:

- *Property models* (Chapter 3). When dealing with 'consumer oriented chemicals based products' (formulations in this work) several product aspects (performances, behaviours) need to be considered. They are all affected by the physical and chemical properties of the mixture/blend. As a result, property models are at the core of product design. Pure compound physical and chemical properties need to be available, at first. Then, the properties of the mixtures/blends need to be determined. Solubility issues are also involved, therefore, models for the estimation of the phase equilibria need to be employed. In this work, existing property models are to be adopted, when model parameters are available and their predictive accuracy is satisfactory. If model parameters are not available, they are to be regressed through the use of experimental data. If the property models are not available, they need to be developed.
- *Method and tools* (Chapter 4). Three new algorithms are to be developed as part of this work: an algorithm for mixture classification, an algorithm for mixture/blend design (MIXD), and an algorithm for evaluation of the phase behaviour of the mixture/blend (STABILITY). The mixture classification algorithm will determine for which mixtures excess properties of mixing are neglectable. The MIXD algorithm will decompose the CAM^{b}D problem into different levels according to the types of models used (linear property models, non-linear property models, phase equilibria models) and then employ mathematical programming techniques in each sub-problem to reduce the number of candidate mixtures/blends. The STABILITY algorithm will screen the solvent mixture according to the trend of the Gibbs energy of mixing, and calculate the solubility limits. Ingredients databases to be used in the algorithms and knowledge base to support some of the design/verification choices are also to be developed.
- *Methodology* (Chapter 5). A methodology for the design and verification of formulated products with a liquid form is to be developed. Due to the complexity of the systems under investigation, this methodology will decompose the systems in sub-problems. In each sub-problem, different tools are to be employed (databases, knowledge base property prediction packages, modelling tool, optimization techniques, mixture design algorithm, stability test algorithm,…).
- *Framework* (Chapter 5). The property models, the methods and tools, and the methodology together with the related databases and knowledge base are to be collected in a systematic framework through which their use will be facilitated

and made more efficient for the design of liquid formulated products. The framework is to be based precisely on the computer-aided stage of the methodology for the design and verification of formulated products, that is, it will include all the tasks and sub-tasks of the methodology. The newly developed framework is to be implemented as part of an existing software, the 'virtual Product-Process Design laboratory' (Morales-Rodriguez, 2009), which hosts dedicated work-flows for the solution of product and process design problems.

- *Case studies* (Chapters 6 and 7). The application of the framework is to be highlighted through a number of case studies involving different types of formulated products.

Figs. 2.2-2.3 summarize the objectives of this PhD project and the interaction/integration between the methodology, the methods and tools, the property models, the databases and the knowledge base for the design and verification of formulated products. Fig. 2.2 shows the integration/interaction for the design scenario, while Fig. 2.3 for the verification scenario.

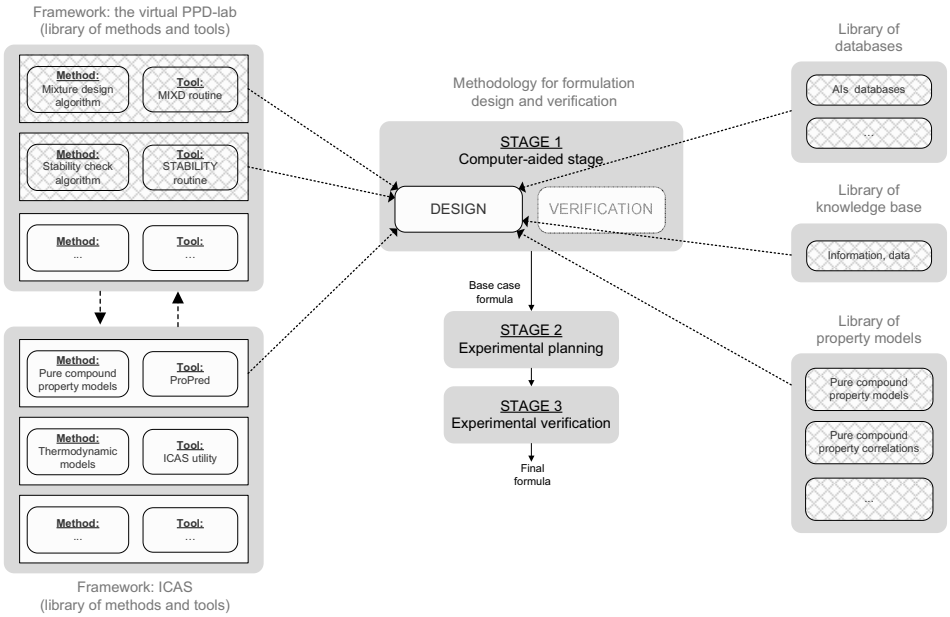


Figure 2.2. Summary of the objectives of this PhD project and their interaction/integration, for the design scenario of the methodology for formulation design and development.

In Figs. 2.2-2.3, the boxes identified by a different pattern represent the methods and tools, the property models, the databases and the knowledge base that are to be

developed in this PhD project. Some of the property models are to be incorporated directly in some of the methods and tools (in MIXD or STABILITY, for instance).
The computer-aided stage of the methodology for formulation design and verification is to be included in the virtual PPD-lab framework, as one of the available work-flows. The methods and tools to be developed for the framework are the MIXD and STABILITY algorithms/computer programs. But also other tools are necessary, such as ProPred, for the pure component property estimation, and ICAS utility, for the generation of various types of phase diagrams employing a big variety of phase equilibria models. These tools are part of another framework (ICAS), but some of the ICAS tools can directly be used from the virtual PPD-lab. ICAS (Nielsen *et al.*, 2001; ICAS Documentation, 2001) is an Integrated Computer Aided System that combines computer-aided tools for modelling, simulation (including property prediction), synthesis/design, control and analysis into a single integrated system.

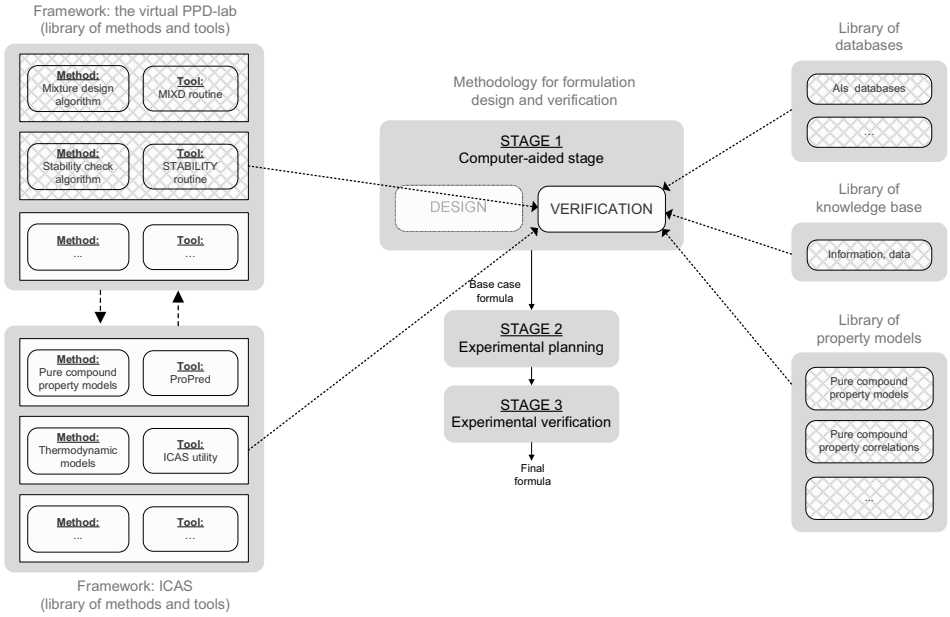


Figure 2.3. Summary of the objectives of this PhD project and their interaction/integration, for the verification scenario of the overall methodology for formulation design and development.

3

TARGET PROPERTY MODELS

Property models are at the core of model-based product design. The reliability of the results obtained in the solution of a product design problem depends very much on the accuracy of the property models employed for the mathematical representation of the chemical systems under consideration. This chapter is dedicated to the estimation of the physicochemical properties employed in this work for the solution of problems related to the design and verification of liquid formulated products.

At first, the physicochemical properties that have to be modelled are highlighted (§3.1). The properties that are necessary are dictated by the performance criteria that have to be satisfied when designing or verifying formulated products. In fact, each set of performance criteria is strictly related to one or more target property. Once the list of necessary properties is known, a search of the open literature is performed, in order to identify the most appropriate models for their estimation. When appropriate models are available, together with their parameters, they are adapted in this work for the necessary property estimation. When the available models are found to be unsatisfactory and/or the model parameters are not available, new models have been developed or new parameters have been obtained for the description of the chemical system/systems under consideration.

3.1 Modelling needs

In this work the physicochemical properties of interest are mixture properties, since formulated products are multicomponent mixtures/blends. But mixture property models need the pure compound property values of the constituent chemicals. Therefore, models for the estimation of both pure compound properties and mixture properties are needed. It has to be underlined that, when experimental data were available, they were directly employed in the calculations.

Table 3.1 shows the set of performance criteria typically considered when designing formulated products (at least the products considered in this work: paint formulation, insect repellent lotion, sunscreen lotion, hair spray). For each performance criteria, the related target properties are also given. The translation of performance criteria into a set of target properties requires insight and experience. In fact, the reliability of the results depends to some extent on this translation. In this work, literature/patent information were employed together with insight gained from the formulated

products literature in order to identify the physicochemical properties related to each performance criteria.

Table 3.1. Performance criteria employed in this work for the design and verification of the products considered (paint formulation, insect repellent lotion, sunscreen lotion, hair spray) and translation of the performance criteria into target properties.

performance criteria	**target properties**
spread-ability	η, σ, ρ
spray-ability	ν, ρ
drying time	T_{90}
flammability	T_f
toxicity	LC_{50}
conductivity	ε
cost	C
solubility	δ / δ_D, δ_P, δ_H / ΔG^{mix}

The meaning of the symbols representing the target properties is explained in Table 3.2, where the models selected (in this work) for the estimation of the mixture target properties are also listed. In the last column, the dependence of the mixture property on the composition, other mixture properties, pure compound properties, temperature and/or pressure is given.

It is evident that pure compound properties are needed in order to estimate the mixture properties. Therefore, models for the estimation of pure compound properties are also necessary. Table 3.3 lists the models employed (in this work) for the estimations of pure compound properties, where the properties are classified in terms of primary and secondary properties. Primary properties depend on the molecular structure, and therefore, they are determined directly from group contribution methods. Secondary properties are not determined (only) by group contribution methods, but are determined by methods that use some of the primary properties as specified variables. For example, the dielectric constant ε_i is a function of the Hildebrand solubility parameter δ_i, and δ_i is a primary property since it is a function of the molecular structure.

The kinematic viscosity, ν, by definition is the ratio of the dynamic viscosity and the density. Hence, no new models are needed for the estimations of this property, provided values for dynamic viscosity and density.

Table 3.2. Target properties and models employed in this work for their estimations. The last column shows the dependence of the target properties on other variables. Highlighted in grey, the models reviewed in this Chapter.

target property	description	model	reference	function
η	dynamic viscosity	linear mixing rule GC(UNIFAC)-based method	definition Cao *et al.*, 1993, §3.4.2	f(η_i, x_i) / f(η_i, x_i, γ_i)
ν	kinematic viscosity	definition	-	f(η, ρ)
σ	surface tension	linear mixing rule GC(UNIFAC)-based method	definition Suarez *et al.*, 1989, §3.4.2	f(σ_i, x_i) / f(σ_i, x_i, γ_i)
ρ	density	linear mixing rule (on the molar volume)	definition	f(ρ_i, x_i)
T_{90}	evaporation time	GC(UNIFAC)-based method	Klein *et al.*, 1992	f($T_{90,i}$, x_i, γ_i)
T_f	open cup flash point	GC(UNIFAC)-based method	Liaw *et al.*, 2002	f($T_{f,i}$, x_i, γ_i)
LC_{50}	toxicity parameter	linear mixing rule	definition	f($LC_{50,i}$, x_i)
ε	dielectric constant	linear mixing rule	definition	f(ε_i, x_i)
C	cost	linear mixing rule	definition	f(C_i, x_i)
δ	Hildebrand solubility parameter	linear mixing rule	definition	f(δ_i, x_i)
δ_D, δ_P, δ_H	Hansen solubility parameters	linear mixing rule	definition	f($\delta_{D,i}$, x_i), f($\delta_{P,i}$, x_i), f($\delta_{H,i}$, x_i)
ΔG^{mix}	delta Gibbs energy of mixing	UNIFAC UNIQUAC NRTL	Fredenslund *et al.*, 1975 Abrams and Prausnitz, 1975 Renon and Prausnitz, 1968	f(GC, segments, x_i)
ΔG^{mix}	for polymer systems	GC-Flory FV-UNIQUAC	Bogdanic and Fredenslund, 1994, §3.5 Bogdanic and Vidal, 2000, §3.6	f(GC, segments, x_i)

GC = Group Contribution (structure of the compound)

Table 3.3. Pure compound properties employed in the models for the estimations of the target properties, and models employed in this work for their estimations. Highlighted in grey, the properties modelled in this work (and presented in this chapter).

pure compound property	type	model	reference	function
η_i	primary	M&G GC$^+$ method	this work, §3.2 (Conte *et al.*, 2008)	f(GC, T)
ν_i	secondary	definition ($\nu_i = \eta_i / \rho_i$)	-	f(η_i, ρ_i)
σ_i	primary	M&G GC$^+$ method	this work, §3.2 (Conte *et al.*, 2008)	f(GC, T)
ρ_i	secondary	modified Rackett correlation	Spencer and Danner, 1972	f($T_{C,i}$, $P_{C,i}$, ω_i, T)
$T_{90,i}$	secondary	correlation	this work, §3.3.1	f(P_i^{sat})
$T_{f,i}$	secondary	C&G GC method	Constantinou and Gani, 1994	f(GC, $T_{b,i}$)
$LC_{50,i}$	primary	GC-based method	Martin and Young, 2001	f(GC)
ε_i	secondary	correlation	Horvath, 1992	f(δ_i) / f(nD_i, Dm_i)
C_i	secondary	correlation	this work, §3.3.2	f(ρ_i)
δ_i	primary	M&G GC method	Marrero and Gani, 2001	f(GC)
$\delta_{D,i}$, $\delta_{P,i}$, $\delta_{H,i}$	primary	M&G GC$^+$ method	Modarresi *et al.*, 2008	f(GC)

The T_{90} evaporation time is the time required for 90% evaporation (by weight) of the pure compound or mixture/blend. The LC_{50} is the aqueous concentration causing 50% mortality in fathead minnows after 96 hours.
In Table 3.3, the property models developed in this work are highlighted. In the sections below, the development of the new models/correlations is reported: in §3.2 the developed M&G GC$^+$ models for the estimation of viscosity and surface tension for pure compounds are presented; in §3.3 the developed correlations for the estimation of the T_{90} evaporation time and the cost for pure compounds are presented.
For the estimation of the mixture properties, no new property models needed to be developed. However, existing models were adapted in this work. §3.4 is dedicated to the models based on linear mixing rules: some considerations about the use of such models for screening purposes are also discussed. In §3.5-3.7 some of the mixture property models listed in Table 3.3 (the one highlighted in dark grey) are reviewed. These models are: the model of Suarez *et al.* (1989) and Cao *et al.* (1993) for the estimation of mixture surface tension and viscosity, respectively (§3.5.1, §3.5.2); the GC-Flory EoS (§3.6) and the FV-UNIQUAC (§3.7).
The mixture property models for surface tension and viscosity are reviewed in order to demonstrate that, with the pure compound property values predicted through the developed M&G GC$^+$ models, the estimation of the mixture properties is feasible. GC-Flory EoS and FV-UNIQUAC are reviewed because they have been extensively employed in Chapter 7, in the hair spray case study.

3.2 Viscosity and surface tension for pure compounds

Surface tension (σ) and viscosity (η) are properties widely used in the design of chemical products and the processes that manufacture them. Knowledge of these properties plays important roles in design issues related to transport of mass and/or energy, wetting, adhesion, friction, spreading, spraying and many more.
Surface tension is a measure of the property of liquids arising from unbalanced molecular cohesive forces at or near the surface, as a result of which the surface tends to contract. While in the bulk of the liquid each molecule may be pulled equally in all directions by neighbouring liquid molecules resulting in a net force of zero, at the surface of the liquid, the molecules may be pulled inwards by other molecules deeper inside the liquid but with no liquid molecules on the outside to balance these forces. Consequently, all the molecules at the surface are subject to an inward force of molecular attraction that can be balanced only by the resistance of the liquid to compression. Hence the liquid squeezes itself together until it has the locally lowest surface area possible.
Viscosity is a property of a fluid that provides a measure of the resistance to flow, that is, it is the fluid resistance to shear or flow and is a measure of the adhesive/cohesive or frictional fluid property. The resistance is caused by intermolecular friction exerted when layers of fluids attempt to slide by one another.

There are two related measures of fluid viscosity: dynamic (or absolute) and kinematic viscosity. Dynamic viscosity is the tangential force per unit area required to move one horizontal plane with respect to the other at unit velocity when maintained a unit distance apart by the fluid. Kinematic viscosity is the ratio between dynamic viscosity and density. In this work, the dynamic viscosity (cP or mPa·s) was considered and unless otherwise indicated, in this chapter the term 'viscosity' will be used to mean 'dynamic viscosity'.

Although some experimental data for pure component surface tension and viscosity can be found (Gani *et al.*, 1991), there is still a large requirement of these data in chemical product-process design. In fact, product design is a combinatorial problem where thousands of candidates are screened and if no data are available for one compound, this has to be taken out from the screening procedure thereby rejecting a potential candidate. Therefore, the development and use of validated models to predict these properties becomes important for applications in product-process design when it is neither practical nor economically feasible to measure them. For the models to be predictive, the use of group-contribution (GC) based methods (Constantinou and Gani R., 1994; Joback and Reid, 1987; Marrero and Gani, 2001) is ideally suited. In theses methods, the molecular structure of an organic chemical is represented by a set of functional groups, where each group contributes in an additive manner to the property or property function under consideration. The GC-based method that has been finding increasing attention is the Marrero and Gani (M&G) method (Marrero and Gani, 2001), which has been employed to predict a wide range of properties covering a very wide range of organic chemicals.

Other approaches are correlations based on the corresponding states theorem, or, quantitative-structure-property relationships, or QSPR (Delgado and Diaz, 2006). The GC approach, however, has been found to provide good results with reasonable accuracy for a wide range of organic chemicals. Also, it is very simple and easy to use.

However, viscosity and surface tension had not been modelled with the M&G method, and therefore, they were modelled in this work. First, the theoretical background is given (§3.2.1): the Marrero and Gani method is described. Then some modelling considerations are discussed (§3.2.2) and in §3.2.3 results are reported. In §3.2.4 the accuracy of the newly developed models is compared with other models.

3.2.1 The Marrero and Gani GC$^+$ models

Like other M&G methods, the estimation of the properties of an organic chemical is performed at three levels (§3.2.1.1). Despite the above mentioned advantages of the GC-based methods, their ranges of applicability are still restricted because of the non availability of the needed experimental data, and therefore, the contributions of the groups. To increase the application range of the GC-based methods, Gani *et al.* (2005a) suggested the creation of missing groups and predicting their contributions

through connectivity indices. When at least one part of the molecular structure cannot be described by available groups and/or their contributions, the GC-based method cannot be used to estimate a property based on the GC approach. For example, the compound shown in Fig. 3.1, needs the group O = PO(O) to represent its molecular structure. But the M&G method does not have this group or its contribution.

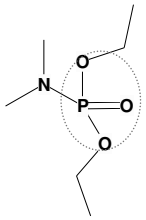


Figure 3.1. Diethyl dimethylphosphoramide (CAS nº 2404-03-7) molecular representation. The red circle represents the missing group in the M&G method: O = PO(O).

The inclusion of this new group in the group set of the M&G method and determination of its contribution from regression is not a practical approach, since this is a lengthy time consuming process, especially when new experimental data are needed to estimate the contribution for the new group with an adequate statistical significance. The methodology proposed by Gani *et al.* (2005a) allows the creation of new groups and the prediction of their contributions for specific chemicals whose atoms and connectivity indices are available in a parameter table of atoms, connectivity indices and their contributions (for the specific property of interest). That is, according to Gani *et al.* (2005a), the contribution of this missing group can be predicted through the atom connectivity index based method.

3.2.1.1 Group-Contribution (GC) based method

In this method, a physicochemical property is predicted considering that each compound can be described by groups at three levels: first-order groups, second-order groups and third-order groups. The (first) basic level uses contributions of first-order groups that describe a wide variety of organic compounds. The higher (second) level provides additional structural information, not provided by the first-order groups and thus corrects the estimates at the first-level. These second-order groups are based on conjugation and they can account for distinctions among a class of isomers but not cis-trans isomers. They are not necessary to represent the total structure of the molecules. A further correction (adjustment) to the prediction is provided through the third-level, where the contributions of parts of the structure of complex molecules are calculated. Like the second-order groups, the third-order groups also use the first-order groups and do not represent the entire molecular structure of the chemical. Note that the second- and third-order groups could be considered as corrections to the

general first-order contributions. Also, the second- and third-order groups use the first-order groups as their building blocks.

If the contributions of these groups are available in group parameter tables, the estimation of surface tension/viscosity is possible in three stages: an initial approximation given by the contribution of first-order groups, an improvement (or correction) provided by the second-order groups, that can be further refined with the third-order groups. Fig. 3.2 illustrates this multilevel approach. The central region representing the first-level contains the contributions of all the first-order groups representing the molecular structure of the chemical (mainly to monofunctional compounds). The (second) middle region includes second-order groups, which are applied mainly to multifunctional complex molecules. The (third) final region includes third-order groups that perform additional corrections to handle the contributions of more complex molecular structures. Properties of large, complex and heterocyclic compounds are handled through these groups.

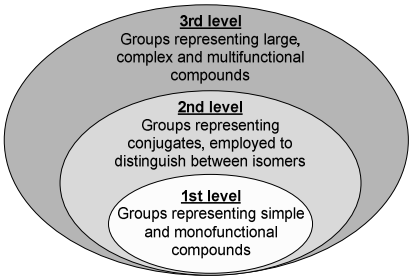


Figure 3.2. Schematic illustration of the multilevel approach for the M&G group contribution method.

The M&G estimation model has the form of the following equation:

$$F(\zeta) = \sum_{i}^{NG1} N_i C_i + y \sum_{i}^{NG2} M_j D_j + h \sum_{i}^{NG3} O_k E_k \qquad (3.1)$$

Where $F(\zeta)$ is a function of the estimated property ζ (surface tension σ or viscosity η); C_i is the contribution of the first-order group of type i; N_i is the occurrence of first-order group i; D_j is the contribution of the second-order group of type j; M_j is the occurrence of second-order group j; E_k is the contribution of the third-order group; O_k is the occurrence of third-order group k; y, h are binary variables. $NG1$, $NG2$ and $NG3$ are the number of first-, second- and third-order groups, respectively.

In the first estimation level, the values of y and h are assigned to zero since, in this level, only first-order groups are involved. In the second level, constants y and h are assigned unity and zero values respectively, because only first and second-order groups are involved. Finally in third level, since all group orders are involved, both constants are equal to one.

The selection of the function $F(\zeta)$ is based on the contributions C_i, D_j, E_k, attempting to:

- Achieve the required addition;
- Exhibit the best possible fit of experimental data;
- Provide good extrapolation capability and therefore, a wide range of applicability.

Usually $F(\zeta)$ includes new adjustable parameters or universal constants.
The contributions of the model C_i, D_j, E_k, are determined through the following regression procedure (Fig. 3.3):

- Determine the contributions C_is of the first-order contribution groups and universal constants; y and h of Eq. 3.1 are set to zero and a regression is carried out;
- Using the estimated values of C_is, the second-order groups are activated in Eq. 3.1 by set the y = 1 and the contributions of second-order groups (D_js) are estimates through regression;
- Finally, to improve the estimation, first- and second- order segments of Eq. 3.1 are considered known (C_i and D_j are fixed in the previous steps), y and h constants are set to the unity and the third-order contribution groups, E_ks, are determined by regression.

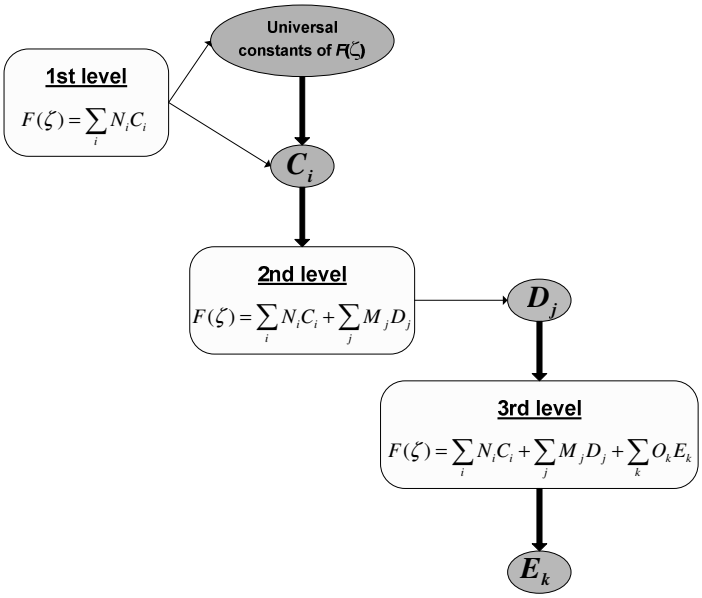


Figure 3.3. The regression procedure to determine the group contributions and the universal constants.

This regression procedure ensures the independence among all three levels of group contributions. Furthermore, the contributions of higher levels serve as corrections to approximations of the lower levels.
The Levenberg-Marquardt method is used for the regression steps where the objective is to minimize the sum of squares of the differences between experimental and estimated values of the property ζ (surface tension or viscosity).
The first-, second- and third-order groups are defined based on the identification criteria described in Marrero and Gani (2001). The M&G group representations for each molecule can be obtained using the ICAS software (Nielsen *et al.*, 2001; ICAS Documentation, 2001).

3.2.1.2 Connectivity Index (CI) based method

The methodology proposed by Gani *et al.* (2005a) permits the creation of missing groups and prediction of their contribution by using valence connectivity indices ${}^{v}\chi$ (Kier and Hall, 1998; Kier and Hall, 2000; Kier and Hall, 2001).
The connectivity indices considered in this work are two:

- ${}^{v}\chi^{0}$ is the zero-order connectivity index, which accounts for the atoms present in the compound;
- ${}^{v}\chi^{1}$ is the first-order connectivity index and accounts for how the atoms are joined together through the chemical bonds.

The connectivity indices are defined via graphical theoretical concepts intended to describe topological characteristics of the molecular structure. This graphical treatment starts by the delineation of the hydrogen-suppressed graph of the molecular structure. Fig. 3.4 gives a representation of a simple molecule (propanoic acid) in terms of its molecular structure and its corresponding hydrogen-suppressed graph.
In this representation, the non-hydrogen atoms become vertices 1, 2, 3, 4 and 5 while the bonds become a, b, c, and d. The omission of hydrogens and double bonds in the graph representation is compensated by the way in which the atomic index δ^{v} for each vertex is defined.

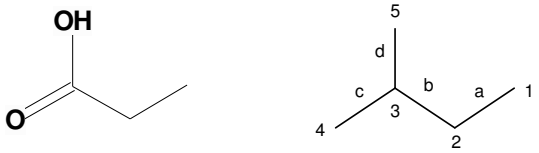


Figure 3.4. Representation of molecular structure (left-side) and hydrogen-suppressed graph for propanoic acid (right side).

Table 3.4 lists the values of atomic indices for various atoms and vertices, where n_H is the number of attached hydrogen atoms.

Table 3.4. Values of atomic index δ^v for each atom/vertex in acyclic, cyclic and special molecules.

atom	δ^v		
	acyclic	**cyclic**	**special**
C	4-n_H	14-n_H	/
Si	4-n_H	14-n_H	/
N	5-n_H	15-n_H	6 (nitro-comp.)
F	7	/	/
Br	7/27	/	/
Cl	7/9	/	/
I	7/47	/	/
Na	1/10	/	/
K	1/18	/	/
O	6-n_H	16-n_H	/
P	5/9	/	1/3 (PH_2)
			4/9 (PH)
S	/	"special"+9	5/9 (SH)
			2/3 (pure S in ring)
			5/9 (S=(different atom))

The bond indices β^k are defined through the pair (that is bonding atoms) of atomic indices δ^v, by the following equation:

$$\beta^k = \delta_i^v . \delta_j^v \tag{3.2}$$

i and *j* are atoms involved in the bond.
The zero-order (atomic) connectivity index ($^v\chi^0$) is defined by:

$$\chi^0 = \sum_{i=1}^{NV} \left(\frac{1}{\sqrt{\delta_i^v}} \right) \tag{3.3}$$

NV is the number of vertices (atoms) in the hydrogen-suppressed graph and the values of δ_i^v can be obtained from Table 3.4.
In the same way, the first-order (bond) connectivity index ($^v\chi^1$) is defined by the summation of the edges of the hydrogen-suppressed graph:

$$\chi^1 = \sum_{i=1}^{NB} \left(\frac{1}{\sqrt{\beta_i^k}} \right) \tag{3.4}$$

NB is the number of bonds in the graph and β_i^k is calculated by Eq. 3.2.
In Table 3.5 calculations are reported for propanoic acid (Fig. 3.4). The atom and bonds indices are calculated using Table 3.4 and Eq. 3.2, respectively. The zero- and first-order connectivity indices are calculated by Eqs. 3.3 and 3.4, respectively.

Table 3.5. Calculated atom and bond indices, first- and second-order connectivity index for propanoic acid.

atoms	**1**	**2**	**3**	**4**	**5**	**bond**	**a (1-2)**	**b (2-3)**	**c (3-4)**	**d (3-5)**
δ_i^v	1	2	4	6	5	β_i^k	2	8	4	20
$1/\sqrt{\delta_i^v}$	1.00	0.71	0.50	0.41	0.45	$1/\sqrt{\beta_i^k}$	0.71	0.35	0.20	0.22
${}^v\chi^0$			3.06			${}^v\chi^1$			1.49	

The same sets of experimental data for pure component surface tension and viscosity employed for the regression of the group contributions for the GC-based method are also used to regress the connectivity indices contributions. But also data corresponding to compounds that cannot be described by the M&G groups can be included in the data set for the atom contributions regression (§3.2.2.3).
In this method, the following model is employed:

$$F(\zeta) = \sum_i^{NA} \left(ac_i AC_i \right) + s\left({}^v\chi^0 \right) + 2t\left({}^v\chi^1 \right) + u \qquad (3.5)$$

$F(\zeta)$ is a function of surface tension or viscosity, the same as Eq. 3.1; ac_i are the occurrences of the i^{th} atoms in the molecular structure; AC_i is the contribution of atom i; s and t are adjustable parameters; u is a constant. *NA* is the number of atoms.

3.2.1.3 Combined GC-CI method

For combined GC-CI method, it is possible to create new groups (missing groups of M&G method or any other host GC-based method) with the CI-based method and then to predict the needed property of the chemical with the host GC-based method. Some rules are however needed to represent groups with connectivity indices. The main rules are represented by the following equations, as proposed by Gani *et al.* (2005a):

$$\left({}^v\chi^0 \right)_{molecule} = \sum_i^{NG} \left({}^v\chi^0 \right)_i \qquad (3.6)$$

$$\left({}^v\chi^1 \right)_{molecule} = \sum_i^{NG} \left({}^v\chi^1 \right)_i \qquad (3.7)$$

i represents a group in the molecule.
Special attention is needed in the determination of $({}^{v}\chi^{1})_i$, due to the fact that the bonds between groups need to be included. Therefore, the following equation is used.

$$\left({}^{v}\chi^{1}\right)_{group} = \sum_{k}^{NBi}\left(\frac{1}{\sqrt{\beta^{k}_{internal\ bonds}}}\right) + \sum_{i}^{NBe}\left(\frac{0.5}{\sqrt{\beta^{m}_{bonds\ out\ of\ groups}}}\right) \tag{3.8}$$

NBi is the number of internal bonds in the group and *NBe* is the number of bonds leaving the group.
For representation of propanoic acid by the M&G method, the first-order groups are: CH_3, CH_2 and COOH. ${}^{v}\chi^{0}$ and ${}^{v}\chi^{1}$ for each group and for the molecule are:

$$\left({}^{v}\chi^{o}\right)_{molecule} = \left({}^{v}\chi^{o}\right)_{CH3} + \left({}^{v}\chi^{o}\right)_{CH2} + \left({}^{v}\chi^{o}\right)_{COOH} =$$
$$= 1.00 + 0.71 + (0.50 + 0.41 + 0.45) =$$
$$= 3.06$$

$$\left({}^{v}\chi^{1}\right)_{molecule} = \left({}^{v}\chi^{1}\right)_{CH3} + \left({}^{v}\chi^{1}\right)_{CH2} + \left({}^{v}\chi^{1}\right)_{COOH} =$$
$$= \left[0.5\cdot(0.71)\right] + \left[0.5\cdot(0.71) + 0.5\cdot(0.35)\right] + \left[0.5\cdot(0.35) + 1\cdot(0.20) + 1\cdot(0.22)\right] =$$
$$= 1.49$$

When applied to groups (missing fragments), the CI-based model (Eq. 3.5) is rewritten in the following form, for each missing fragment *k*:

$$F\left(\zeta\right)_k = \sum_{i}^{NA}\left(ac_i AC_i\right) + s\left({}^{v}\chi^{0}\right) + 2t\left({}^{v}\chi^{1}\right) \tag{3.9}$$

Then, all fragments are combined together:

$$F\left(\zeta^{*}\right) = \sum_{k}^{NK}\left(\upsilon_k F\left(\zeta\right)_k\right) + u \tag{3.10}$$

$F(\zeta^{*})$ is a function of surface tension or viscosity for all missing groups/fragments; $F(\zeta)_k$ is a function of surface tension or viscosity for missing group/fragment *k*; *NK* is the number of missing groups/fragments; υ_k is the number of times a missing group/fragment *k* appears in the molecule.
The final model is obtained by the combination of GC-CI methods, and it is described by:

$$F(\zeta) = \sum_{i}^{NG1} N_i C_i + F\left(\zeta^{*}\right) + y\sum_{i}^{NG2} M_j D_j + h\sum_{i}^{NG3} O_k E_k \tag{3.11}$$

The combined GC-CI method, where the GC-based method is the M&G method, is also referred to as the M&G GC$^+$ method. The prediction of surface tension (or viscosity) is highlighted through the step-by-step procedure in Fig. 3.5.

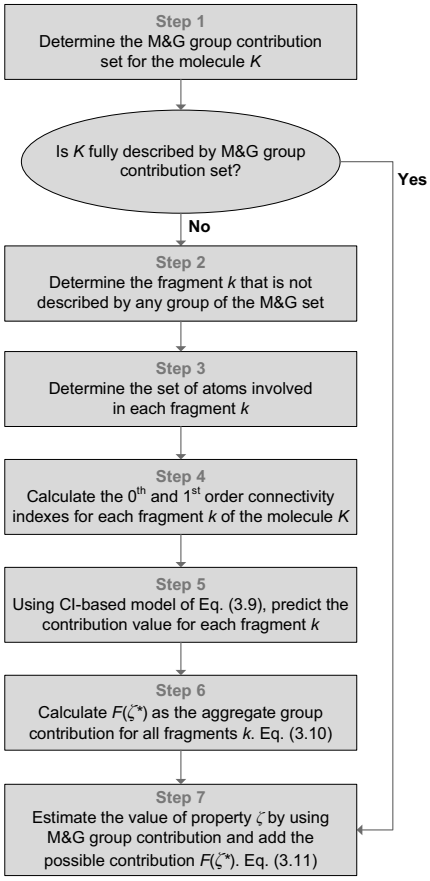


Figure 3.5. Step by step procedure to predict a general physicochemical property of pure compounds using combined GC-CI method (M&G GC$^+$).

3.2.2 Modelling considerations

In order to develop the M&G GC$^+$ methods three main actions need to be performed:

- Collect the experimental data to employ for the GC and CI parameter regressions (§3.2.2.1);
- Choose the property (additive) function $F(\zeta)$ (§3.2.2.2);
- Regress the group contributions and the atom contributions using the collected experimental data (§3.2.2.3).

3.2.2.1 Experimental data

A data set of pure component surface tension of organic chemicals (at 298 K and atmospheric pressure) was collected from different sources (Dean, 1992; Yaws, 1999). The reported experimental values of surface tension for a total of 420 structurally diverse, including complex and polyfunctional molecules, organic chemicals were selected to develop the model. Of these 420 compounds, 2 are not described by the M&G groups set. Appendix A lists all the data points.
For viscosity, another data set of pure component property values was collected mainly from the book of Viswanath and Natarajan (1989) and the book of Weast (1984). A set of 453 compounds was selected, including different types of structures as for surface tension. Of these 453 compounds, 8 are not described by the M&G groups set. Appendix B lists the experimental data points together with their sources.

3.2.2.2 Selection of property function

An important step in the M&G model is the choice of the property function $F(\zeta)$. The selection of this function is based on the following criteria:

- The function has to achieve additivity in the contributions of C_i, D_j and E_k;
- It has to exhibit the best possible fit of the experimental data;
- It should provide good extrapolating capability and therefore, a wide range of applicability.

In order to identify $F(\zeta)$, experimental data have been plotted against the occurrence of the CH_2 group for well-known families of compounds, for surface tension (Fig. 3.6a and b) and viscosity (Fig. 3.6c and d).
Note that in the viscosity plots (3.6c and d) the y-axis is chosen as the natural logarithm of the viscosity. It can be assumed that the trends in both plots are linear.
Hence, for the surface tension and viscosity Eq. 3.11 becomes, respectively:

$$\sigma = \sum_{i}^{NG1} N_i C_i + F\left(\sigma^*\right) + y \sum_{i}^{NG2} M_j D_j + h \sum_{i}^{NG3} O_k E_k \tag{3.12}$$

$$\ln\left(\eta\right) = \sum_{i}^{NG1} N_i C_i + F\left(\eta^*\right) + y \sum_{i}^{NG2} M_j D_j + h \sum_{i}^{NG3} O_k E_k \tag{3.13}$$

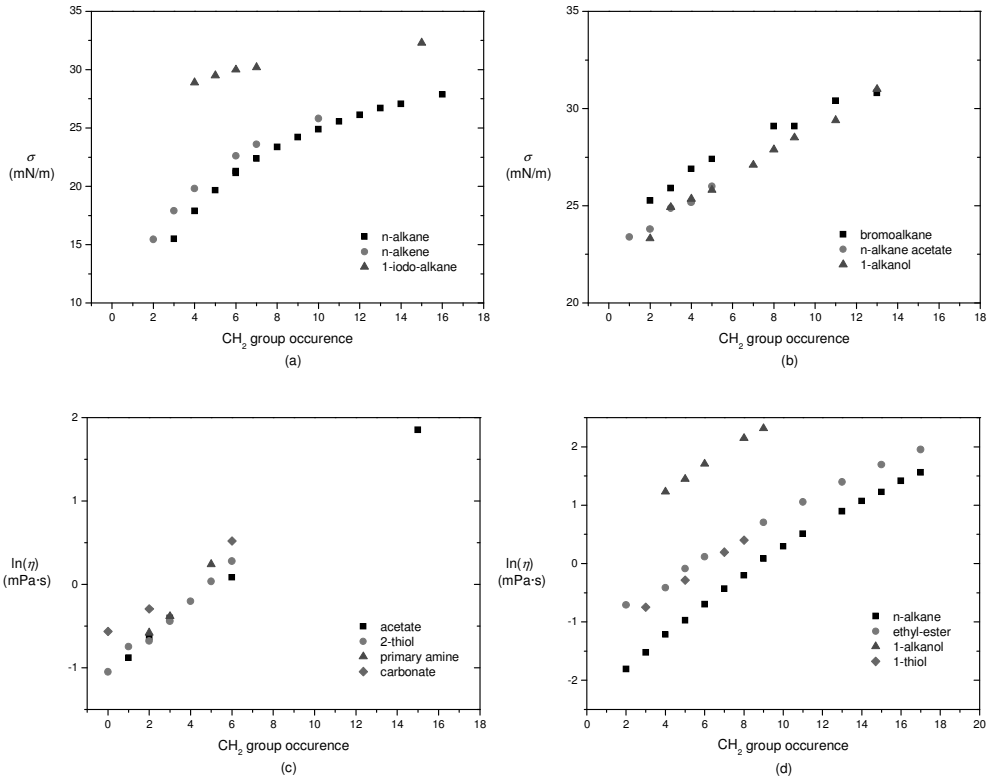


Figure 3.6. (a)-(b) Surface tension at 298 K versus the occurrence of the group CH_2 for different families of compounds. (c)-(d) Natural logarithm of the viscosity at 300 K versus the occurrence of the group CH_2 for different families of compounds.

3.2.2.3 Parameter regression

Not all the data in the open literature can be trusted. Reason for inaccurate experimental measurements can be many, for example, a bad temperature control, or a bad calibration of the instrument. It is obvious that these values could lead the parameter regression to a wrong direction. In order to avoid this problem, it is necessary to identify outliers that could disturb the parameter regression. After the outliers have been identified, the parameter regression can be carried out.

Fig. 3.7 shows the methodology employed in the parameter regression. At first, outliers are identified, then, the parameter regression is carried out for GC-based method, and for the CI-based method. Note that one advantage of this procedure is that the same data set is used to develop two models and then one model (CI-based method) is used to extend the application range of the other (GC-based method).

The procedure is divided into 4 levels:

- Level 1: division in families. The compounds included in the original data set (excluding the compounds that cannot be described with the M&G groups) were divided in terms of different sets (families), according to the type of atoms present in the molecule. Four main families were considered: compounds with only carbon and hydrogen atoms (C,H); compounds with carbon, hydrogen and oxygen atoms (C,H,O); compounds with carbon, hydrogen, oxygen atoms and one halogen (C,H,O,1Hal); all the other compounds (Others) that are not part of the above families. This last set includes also multifunctional molecules. The family classification with examples is given in Table 3.6.
- Level 2: outliers identification.
- Step L2-1. The (C,H) family of compounds was employed for the first parameter regression with the M&G GC-based method (Eq. 3.1). The outliers were identified as those compounds that show a very high error and that did not fit in the average trend of the other compounds. After the outliers were removed (from the original data set) a new (C,H) family (outliers-free) was available and the parameters were regressed again on this new data set;
- Step L2-2. The parameter values (C_i, D_j, E_k) obtained at step L2-1 on the outliers-free data set were kept fixed, and the (C,H,O) family was employed for the second parameter regression. Outliers were identified and removed from the data set, and the parameters were regressed again on the new (C,H,O) family;
- Step L2-3. All the parameters (C_i, D_j, E_k) obtained at step L2-1 and L2-2 were kept fixed, and the (C,H,O,1Hal) family was used for another parameter regression. Outliers were identified and removed from the data set, and the parameters were regressed again on the new (C,H,O,1Hal) family;
- Step L2-4. The remaining compounds (Others), mainly multifunctional compounds, were employed for the final regression, while all the parameters (C_i, D_j, E_k) obtained from step L2-1 to step L2-3 were kept fixed. The outliers were identified and removed from the data set, and the parameters were regressed again on the new (Others) family.
- Level 3: all the outliers identified from steps L2-1 to L2-4 were removed from the original data set. An outliers-free data set was now available, and it was employed to regress the GC model parameters all at once. The parameters values C_i, D_j, E_k obtained in level 2 were used as initial estimates.

- Level 4: the outliers-free data set plus the list of compounds that could be described with the M&G group set (surface tension: 2 compounds; viscosity: 8 compounds) was employed for the parameters regression of the CI-based model.

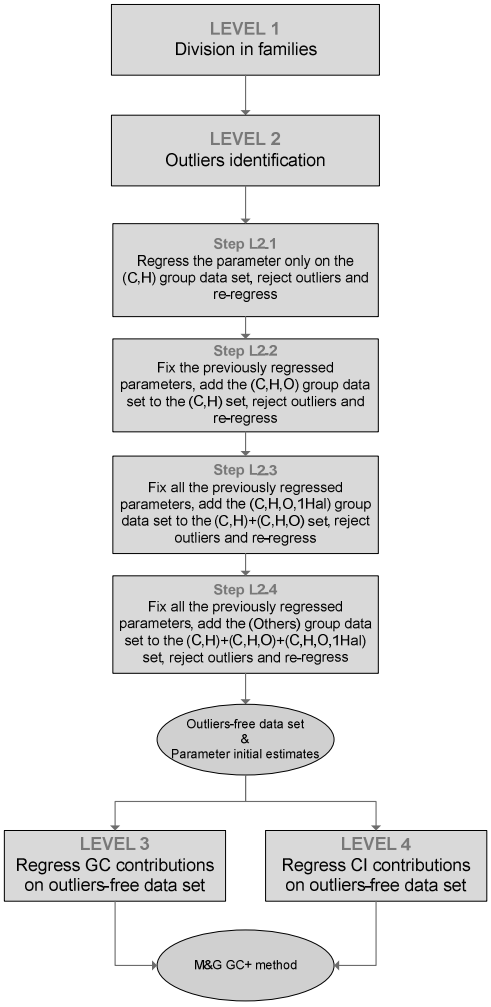


Figure 3.7. Flowchart of the methodology employed for the parameter regression.

A total of 16 outliers were found for the surface tension model, while a total of 15 outliers were found for the viscosity model (see Table 3.7 for more details).

Table 3.6. Families of compounds considered in the methodology for the parameter regression.

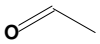
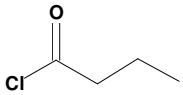
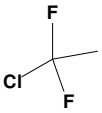
Families	Description	Example	
C,H	only carbon and hydrogen atoms		n-butane
C,H,O	carbon, hydrogen and oxygen atoms		acetaldehyde
C,H,O,1Hal	carbon, hydrogen, oxygen atoms and one halogen atom		butanoyl chloride
Others	all compounds not included in the previous groups, mainly multifunctional molecules		1-chloro-1,1-difluoro-ethane,

Table 3.7. Summary of the number of data collected, the outliers, and the data considered in the GC and the CI parameter regression.

	Surface tension	Viscosity
Total data collected	420	453
Data not described by M&G groups	2	8
Data described by M&G groups	420-2=418	453-8=445
Outliers	16	15
Data considered for the GC regression	418-16=402	445-15=430
Data considered for the CI regression	402+2=404	430+8=438

3.2.3 Results and discussion

The estimated values of surface tension and viscosity, respectively, for the set of compounds employed in the parameter regression are reported in Appendices A and B, respectively. The total list of groups and their regressed contribution values are listed in Appendix C. The application of the M&G method and its ability to predict and distinguish between some isomers is shown through case studies given in Appendix D.
In this section, only statistics are reported.

3.2.3.1 GC-based method

The Standard Deviation (*SD*) and the Average Absolute Error (*AAD*) are defined by Eq. 3.14-3.15, respectively:

$$SD = \sqrt{\frac{\sum_i \left(\zeta_i^{est} - \zeta_i^{exp}\right)^2}{N^{tot}}} \qquad i = 1 \text{ to } 402 \text{ or } 430 \qquad (3.14)$$

$$AAD = \frac{\sum_i \left|\zeta_i^{est} - \zeta_i^{exp}\right|}{N^{tot}} \qquad i = 1 \text{ to } 402 \text{ or } 430 \qquad (3.15)$$

ζ_i^{est} is the surface tension (or viscosity) estimated by the regression and ζ_i^{exp} is the experimental value, for compound *i*.
The correlation statistics for surface tension and viscosity are given in Table 3.8, where the total numbers of data points for each property are reported for each level of group contribution with the corresponding standard deviations.

Table 3.8. Comparison of the standard deviation *SD* for the first-, second- and third-order group contributions.

Surface tension				Viscosity			
Data nº	1st	2nd	3rd	Data nº	1st	2nd	3rd
402	2.04	-	-	430	3.44	-	-
163	-	1.30	-	181	-	1.05	-
9	-	-	0.01	11	-	-	1.64

Note that all the compounds are described by first-order groups (402 compounds for surface tension and 430 for viscosity), but only some of the compounds are described by the second-order groups (163 compounds for surface tension and 181 for viscosity), and even less by the third-order groups (9 compounds for surface tension and 11 for viscosity).
From Table 3.8 it can be noted that a general reduction of the standard deviations has been achieved. Note also that for viscosity correlation, some of the 11 compounds that needed third-order groups, do not need the second-order groups, and therefore, they reflect a reduction of the standard deviation with respect to the first-order standard deviation.
In Table 3.9 the Standard Deviation (*SD*) and the Average Absolute Deviation (*AAD*) are given for all three levels separately. The results showed for second- and third-orders includes all the compounds (N^{tot} = 402 for surface tension; N^{tot} = 430 for viscosity), even those that do not have second- and third-order group contributions. Consequently, the *SD* (or *AAD*) of the third level corresponds to the global result of the three-level group contribution approach.

Table 3.9. Comparison of deviations considering all the compounds and the contributions of each level.

	Surface tension			Viscosity		
	1st	1st and 2nd	1st, 2nd and 3rd	1st	1st and 2nd	1st, 2nd and 3rd
SD	2.04	1.61	1.47	3.44	1.05	0.89
AAD	1.31	1.11	1.05	0.65	0.43	0.37

Plots of the estimated values of surface tension and viscosity with the new GC-based methods are compared with experimental data in Fig. 3.8a and Fig. 3.8b, respectively. For surface tension, the squared correlation coefficient has the value 0.959 while for viscosity, it is 0.985, indicating for both cases, a fine correlation of the data.

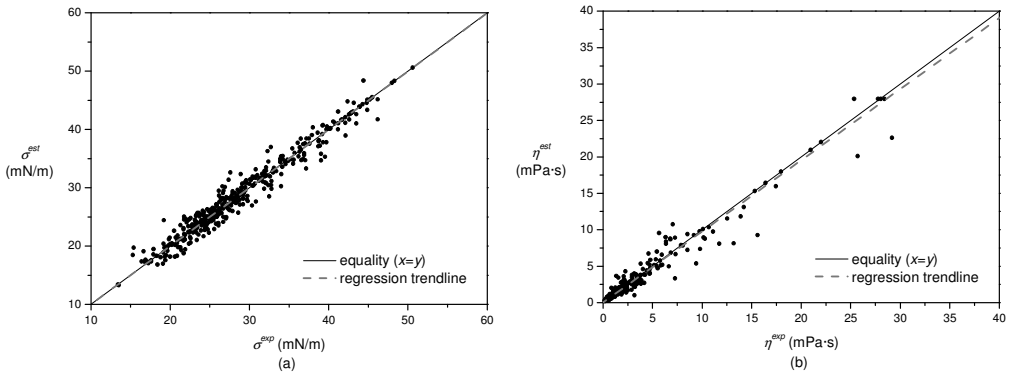


Figure 3.8. (a) Predicted surface tension (σ^{est}) versus experimental data (σ^{exp}) and (b) predicted viscosity (η^{est}) versus experimental data (η^{exp}) for all compounds used for the GC parameter regression, using the GC-based method.

The regression obtained (regression trend lines), for the surface tension model and the viscosity model are:

$$\sigma^{est}\left(\text{mN/m}\right) = 0.9993\,\sigma^{exp} \tag{3.16}$$

$$\eta^{est}\left(\text{mPa}\cdot\text{s}\right) = 0.9768\,\eta^{exp} \tag{3.17}$$

3.2.3.2 CI-based method

Once the function for the property being modelled has been selected, it is possible to estimate the parameters (*s, t, u, ac_i*) of the CI-based model (see Eq. 3.5) using the same data set of experimental data used for the GC-based method (the outlier-free data set) plus the data for those compounds not described by the M&G groups (see Table 3.7).
The regressed CI-based method parameters are listed in Table 3.10.

Table 3.10. Regressed parameters for CI-based method, for both surface tension and viscosity.

Atom	Surface Tension	Viscosity
ac(H)	1.12	0.04
ac(Cl)	11.24	0.46
ac(Br)	20.08	1.24
ac(F)	3.21	0.10
ac(I)	26.26	0.95
ac(N)	13.37	0.53
ac(O)	6.68	0.49
ac(P)	15.76	-0.21
ac(S)	19.00	0.33
ac(C)	2.21	0.31
ac(Si)	-15.71	0.00
s	1.13	-1.01
2*t*	-7.34	0.87
u	0.00	0.00

The use of the values listed in Table 3.10 is highlighted for diethyl dimethylphosphoramidate (Fig. 3.1), assuming that one group/fragment (O = P(O)O) is missing. The contribution for the surface tension is:

$$F(\sigma)_{O=P(O)O} = (3 \cdot AC_O + 1 \cdot AC_P) + s\left({}^{v}\chi^{0}\right)_{O=P(O)O} + 2t\left({}^{v}\chi^{1}\right)_{O=P(O)O} =$$
$$= (3 \cdot 6.68 + 1 \cdot 15.76) + 1.13 \cdot 2.57 - 7.34 \cdot 4.15 = 8.24$$

$$F(\sigma^{*}) = F(\sigma)_{O=P(O)O} + u = 8.24$$

While for viscosity the contribution is:

$$F(\eta)_{O=P(O)O} = (3 \cdot AC_O + 1 \cdot AC_P) + s\left({}^{v}\chi^{0}\right)_{O=P(O)O} + 2t\left({}^{v}\chi^{1}\right)_{O=P(O)O} =$$
$$= \left[3 \cdot 0.49 + 1 \cdot (-0.21)\right] - 1.01 \cdot 2.57 + 0.87 \cdot 4.15 = 2.28$$

$$F(\eta^{*}) = F(\eta)_{O=P(O)O} + u = 2.28$$

Table 3.11 gives a summary of the regression statistics in terms of *SD* and *AAD* for both methods used alone (GC-based and CI-based methods) for each property and their respective datasets.

Fig. 3.9 highlights the comparison of the estimated and experimental values for the two properties.

The results of Table 3.11 and Fig. 3.9, confirm that while the CI-based method covers a wider range of compounds, it cannot be expected to have the same level of accuracy (with a smaller number of parameters) as the corresponding host GC-based method.

The CI-based method alone does not provide improvement to the estimation of the surface tension or the viscosity. However, the CI-based method still gives sufficiently good results, and therefore, can be used as an auxiliary tool for the host GC-based method (the M&G method) to predict the missing group contributions. Since usually not more than 1-2 groups are expected to be missing for any molecular structure of an organic chemical, the prediction of the property would be of acceptable accuracy with the combined GC-CI based method.

Table 3.11. Comparison between GC and CI-based methods estimations for surface tension and viscosity.

	Surface tension		**Viscosity**	
	GC	**CI**	**GC**	**CI**
SD	1.47	6.01	0.89	6.89
AAD	1.05	4.60	0.37	2.10
nº compounds	402	404	430	438

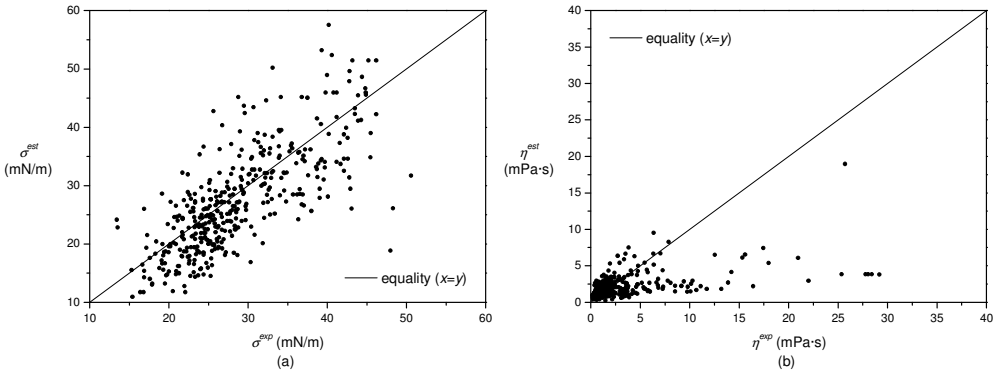


Figure 3.9. (a) Predicted surface tension (σ^{est}) versus experimental data (σ^{exp}) and (b) predicted viscosity (η^{est}) versus experimental data (η^{exp}) for all compounds, using only the CI-based method.

3.2.3.3 Predictive capability

For the viscosity model, a total of 101 extra data points (not included in the parameter regression) were collected. Fig. 3.10 highlights the comparison of the predicted and experimental values. The standard deviation for this set of 101 data points is 3.07.
Appendix E reports the dataset employed for the model testing, the experimental values and the viscosity estimated values.

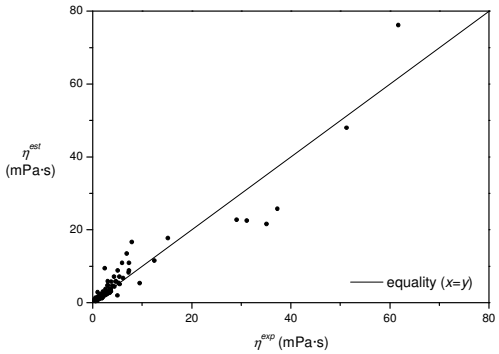


Figure 3.10. Predicted viscosity (η^{est}) versus experimental data (η^{exp}) for a set of compounds not included in the parameter regression.

Table 3.12 highlights the performance of the viscosity M&G method for 5 of the 101 compounds on which the model was tested (5 of the compounds in Fig. 3.10). The predictive performance is acceptable.
The % Relative Deviation (*RD*) corresponds to:

$$RD(\%) = \frac{\left|\zeta_i^{exp} - \zeta_i^{est}\right|}{\zeta_i^{exp}} \cdot 100 \tag{3.18}$$

Table 3.12. Relative Deviation *RD* (%) of M&G method for five compounds not considered in the parameter regression.

Compound	CAS nº	σ (mN/m)			η (mPa·s)		
		exp	est	*RD%*	exp	est	*RD%*
ethyl benzoate	43-89-0	34.50	33.91	1.72	2.01	2.00	0.36
4-methyl-2-pentanol	108-11-2	22.60	22.66	0.27	2.24	2.40	6.92
2-hexanol	626-93-7	23.60	23.78	0.75	3.10	2.70	12.84
m-methylbenzenamine	108-44-1	37.90	39.02	2.96	3.03	2.83	6.69
3-bromoaniline	591-19-5	47.70	45.35	4.92	5.46	5.07	7.15

3.2.4 Comparison with other models

The developed M&G GC$^+$ model for the prediction of viscosity was compared with other existing models. The models considered for comparison were those of Sastri and Rao (1992) and Orrick and Erbar (1974), which have widely been used in the past for the prediction of viscosity and surface tension, respectively. Sastri and Rao (1992) proposed the following equation for the dynamic viscosity η (mPa·s):

$$\eta = \eta_B \left(P^{sat}\right)^{-N_{S\&R}} \tag{3.19}$$

η_B and $N_{S\&R}$ are parameters calculated with a group contribution based model:

$$\eta_B = \sum_i^{NG} \Delta\eta_{B,i} + \sum_i^{NG} \Delta\eta_{cor,i} \tag{3.20}$$

$$N^{S\&R} = 0.2 + \sum_i^{NG} \Delta N_i^{S\&R} + \sum_i^{NG} \Delta N_{cor,i} \tag{3.21}$$

$\Delta\eta_B$, $\Delta\eta_{cor}$, $\Delta N^{S\&R}$ and ΔN_{cor} are group contribution parameters from the Sastri and Rao (1992) parameter table. P^{sat} (atm) is the vapour pressure to be calculated with the following equation:

$$ln\, P^{sat} = \left[4.54 + 1.03 \cdot ln\left(T_b\right)\right] \cdot \left[1 - \frac{3 - 2\left(T/T_b\right)}{T/T} - 0.38\left(3 - 2\left(T/T_b\right)\right)^{0.19} \cdot ln\left(T/T\right)\right] \tag{3.22}$$

T and T_b (K) are the temperature at which the viscosity has to be calculated and the boiling point, respectively. Eq. 3.22 can be used only when $T < T_b$. This correlation is not the most accurate to evaluate the vapour pressure but the model parameters (group contributions) of the Sastri and Rao model have been regressed using Eq. 3.22, so it is recommended to employ this equation for evaluating the vapour pressure.

Orrick and Erbar (1974) model evaluates the viscosity as:

$$ln\left(\frac{\eta}{\rho \cdot Mw}\right) = A^{O\&E} + \frac{B^{O\&E}}{T} \tag{3.23}$$

ρ (g/cm^3) is the liquid density at 293 K (20 °C), Mw (g/mol) is the molecular weight, while $A^{O\&E}$ and $B^{O\&E}$ are taken from the Orrick and Erbar (1974) parameter table.

The viscosity of 12 compounds was calculated with the M&G GC$^+$ model (this work), Sastri and Rao (1992) model and Orrick and Erbar (1974) model. The results are reported in Table 3.13.

The squared residual R^2 is defined as:

$$R^2 = \left(\zeta_i^{est} - \zeta_i^{exp}\right)^2 \tag{3.24}$$

with ζ_i^{est} and ζ_i^{exp} defined as in Eqs. 3.14 and 3.15.

In the last column, the one reporting the errors for Orrick and Erbar method (1974), some values are missing since the parameter table does not contain the groups to describe those molecules. In fact, Orrick and Erbar (1974) method cannot be used with compounds containing nitrogen, sulphur and fluorine. In addition, this method gives high errors, while Sastri and Rao (1992) method can predict the viscosity quite well, but the M&G method shows a far better performance in evaluating the viscosity of pure liquids.

Table 3.13. Squared residuals R^2 on the viscosity calculated with the Marrero and Gani (M&G), Sastri and Rao (S&R) and Orrick and Erbar (O&E) methods; in the last row the standard deviation *SD* is reported.

Compound	Group	R^2		
		M&G	S&R	O&E
nonadecane	alkane	0.00	8.18	400.70
1-heptene	alkene	0.00	0.01	0.00
benzyl alcohol	alcohol	0.00	18.63	3.37
butyl acetate	ester	0.00	0.10	0.00
ethyl tert-butyl ether	ether	0.00	0.02	0.00
benzylamine	amine	0.00	0.00	-
2-octanethiol	thiol	0.00	8.78	-
3-nitro-2-hexene	N-compound	0.00	0.00	-
3-bromotoluene	Br-compound	0.00	0.07	0.24
2-chloro-6-methylpropane	Cl-compound	0.00	0.00	0.00
2-fluorotoluene	F-compound	0.00	0.13	-
chlorodifluoromethane	freon	0.00	0.02	-
SD		0.02	1.73	7.60

In conclusion, after evaluation of Table 3.13, it can be stated that the M&G method or the prediction of viscosity of liquids developed in this work is far superior to the other methods present in the literatre (Sastri and Rao, 1992; Orrick and Erbar, 1974).

3.3 Evaporation time and cost for pure compounds

Property values for the T_{90} evaporation time and the cost of pure compounds are scarce in the literature. To overcome this problem, simple correlations were developed in this work.

3.3.1 Evaporation time

Van Wesenbeeck *et al.* (2008) proposed the following correlation between vapour pressure (P^{sat}) and evaporation rate (*ER*):

$$ln(ER) = 0.865 \cdot ln(P^{sat}) + 12.70 \qquad (3.25)$$

The units of measure for *ER* and P^{sat} are ($g \cdot cm^{-2} \cdot s^{-1}$) and (Pa), respectively.
This correlation describes the evaporation rate quite accurately, in fact the sum of square residuals corresponds to 0.93 (Van Wesenbeeck *et al.*, 2008).
Other correlations have been proposed in the literature (Woodrow *et al.*, 2001), and all show a quite good accuracy in the description of the relation *ER*-P^{sat} ($\Sigma R^2 = 0.99$).
The evaporation rate is the speed with which a solvent evaporates, but the property that is needed in this work is not the evaporation rate, but the T_{90} evaporation time, which is the time needed for 90% (by weight) of the solvent to evaporate. In fact, the property model for the calculation of the mixture evaporation time requires the pure compound T_{90} values.

If the evaporation rate *ER* is a function of the vapour pressure, also the T_{90} should be a function of the vapour pressure. The available T_{90} experimental data (Klein *et al.*, 1992) are plotted versus the vapour pressure in Fig. 3.11. In the same plot, a linear correlation of the data is also shown, which corresponds to Eq. 3.26.

$$ln(T_{90}) = -0.793 \cdot ln(P^{sat}) + 12.416 \tag{3.26}$$

The units of measure for T_{90} and P^{sat} are (s) and (Pa), respectively.

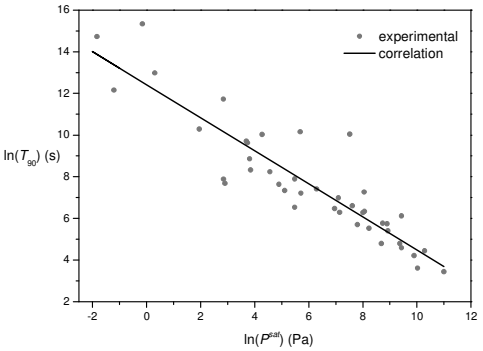


Figure 3.11. Logarithm of experimental T_{90} data versus the logarithm of the vapour pressure. The correlation of the data (Eq. 3.26) is also shown.

The sum of square residuals corresponds to 0.83.
The T_{90} of compounds for which no experimental data are available can now be estimated through Eq. 3.26.

3.3.2 Cost

The correlation of cost data is more difficult than the correlation of evaporation time data. Cost is not a chemical or physical property, therefore it is not straightforward to identify a dependency on other chemical properties or on the molecular structure. In addition, the cost data are really scarce (www.icis.com).
Fig. 3.12 shows the costs (*C*) of compounds belonging to different chemical families (esters, alcohols) versus the molar volume (*V*). A linear dependence of the cost of compounds belonging to the same family can be noted.
The linear correlations of Fig. 3.12 correspond to Eq. 3.27 for the alcohols, and Eq. 3.28 for the esters.

$$C = 2.152 \cdot V - 38.714 \tag{3.27}$$

$$C = 2.356 \cdot V - 119.000 \tag{3.28}$$

The units of measure for the cost and the molar volume are ($/kmol) and (l/kmol), respectively. The sum of square residuals for the alcohols is 0.89 and for the esters is 1.00 (but the data fitted are really few).

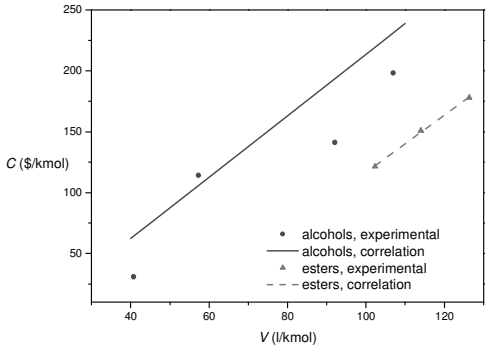


Figure 3.12. Values of cost versus molar volume for some alcohols and esters for which cost data were available.

The cost of alcohols and esters (solvents employed in the case studies developed in this work) for which cost data are not available can now be estimated with Eqs. 3.28 and 3.29, respectively.
It has to be underlined that this model does not take into account the fluctuation of the market, and it has been developed only to fill the lack of chemical prices in the literature.

3.4 Linear models of mixing

A linear mixing rule corresponds to the linear combination of the properties of the pure compounds, weighted in terms of composition of the chemicals in the mixture. For the generic mixture property ζ, the mixture property model based on a linear mixing rule is:

$$\zeta = \sum_{i=1}^{NC} x_i \cdot \zeta_i \tag{3.29}$$

x_i is the composition of compound i in the mixture, and ζ_i the pure compound property.
Linear property models give good predictions of mixture properties for chemical systems that have negligible excess properties of mixing. For chemical systems having large excess property values, the linear models cannot be used. Hence, it could be very useful to determine if mixtures are likely to have negligible excess properties of mixing. Thereby, time consuming and tedious calculations (with rigorous models accounting for excess properties of mixing) can be avoided.

Based on the above discussion, the concept of classifying chemical systems (mixtures) in terms of their hydrogen bonding property, has been considered in this work. Organic chemicals can be classified according to hydrogen bonding (HB) into 3 kinds of fluid: normal (NF), such as alkanes; polar non-associating (PNA), such as esters; and polar associating (PAS), such as alcohols, glycols, etc. If these fluids are combined in binary mixtures, six types of mixtures are possible: NF/NF, PNA/PNA, PAS/PAS, NF/PNA, NF/PAS, PNA/PAS. In Fig. 3.13 (Smith, Van Ness and Abbott, 2001) the excess properties of more than 130 binary mixtures of solvents are shown.

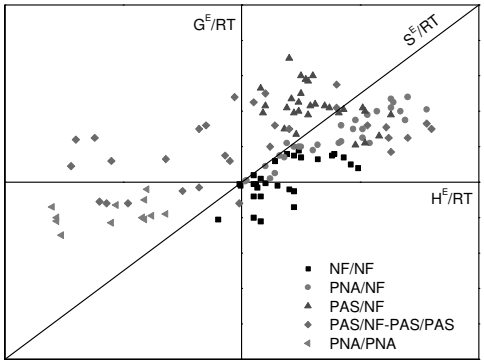


Figure 3.13. Equimolar excess properties for more than 130 mixtures at 298.15 K. G^E, H^E and S^E are the excess Gibbs energy, the excess enthalpy and the excess entropy, respectively. The mixtures are represented according to the hydrogen bonding classification (Smith, Van Ness and Abbott, 2001).

It can be noted that mixtures of two similar fluids, especially of two normal fluids, fall in the region where the excess properties are negligible. Hence, for these kinds of mixtures the use of linear models for the calculation of mixture properties would most likely give reliable predictions.
The following two rules are proposed:

- For mixtures of similar fluids such as NF/NF and PNA/PNA (but not PAS/PAS), linear mixing rules give reliable results since the excess properties of mixing do not show high contributions;
- For mixtures of the other kinds, the excess property values are not negligible. Linear mixing rules do not give reliable results and more reliable mixture property models need to be employed.

3.5 Mixture viscosity and surface tension

The model developed by Suarez *et al.* (1989) for the prediction of the surface tension of mixtures and the model of Cao *et al.* (1993) for the prediction of mixture viscosity can be employed as an alternative to the linear mixing rule. Since these models are quite complex (the surface tension model requires iteration), they cannot be employed

as models for screening of alternative designs. They should be used instead, only for the property calculation for those mixtures that have been identified to have significant excess property values.

3.5.1 Surface tension

The model developed by Suarez *et al.* (1989) is based on the assumption that the surface layer can be treated as a separate phase located between the vapour and the bulk phases. The model equation is:

$$\sigma = \sigma_i + \frac{RT}{A_i} ln\left(\frac{x_{i,s} \cdot \gamma_{i,s}}{x_{i,b} \cdot \gamma_{i,b}}\right) \tag{3.30}$$

σ (mN/m) is the surface tension of the mixture; σ_i (mN/m) is the surface tension of the pure compound i; A_i is the molar surface area of pure component i; $x_{i,s}$ and $x_{i,b}$ are the composition of compound i in the surface and bulk liquid phase, respectively; $\gamma_{i,s}$ and $\gamma_{i,b}$ are the activity coefficients of compound i in the surface and bulk liquid phase, respectively. T is the temperature and R the universal constant of gases.
At the same time, the following equality has to be satisfied:

$$\sum_{i=1}^{NC} x_{i,s} = 1 \tag{3.31}$$

NC is the number of compounds involved in the mixture.
Eqs. 3.30 and 3.31 represent NC+1 equations with NC+1 unknowns (σ and NC values of $x_{i,s}$). The equations are solved by means of the Newton-Raphson method.
The activity coefficients $\gamma_{i,s}$ and $\gamma_{i,b}$ are calculated with the original UNIFAC method (Fredenslund *et al.*, 1975). The molar surface areas A_i can be calculated in two different ways: using the UNIFAC area parameter Q_k (which is calculated from the Van der Waals surface area with the GC-based method given by Bondi, 1968) as shown in Eq. 3.32 or the Paquette (Suarez *et al.*, 1989) equation as shown in Eq. 3.33:

$$A_i = 2.5 \cdot 10^9 \sum_k \upsilon_{k,i} Q_k \tag{3.32}$$

$$A_i = 1.021 \cdot 10^8 V_{c,i}^{\ 6/15} \cdot V_i^{4/15} \tag{3.33}$$

$\upsilon_{i,k}$ is the number of groups of type k in the molecule i; $V_{c,i}$ and V_i (cm^3/mol) the critical volume and the bulk liquid molar volume of compound i, respectively.
The pure compound surface tensions σ_i can be calculated with the M&G GC$^+$ model discussed in §3.2.
Table 3.14 shows the results for the mixture surface tension prediction. The predictions are at a temperature of 298 K. Since for every mixture the surface tension predictions have been carried on for different compositions (the composition range is

also shown) the standard deviation for every mixture is the average on the different compositions; in the last row the global average of the standard deviations is reported.

Table 3.14. Standard deviation on the mixture surface tension predictions for some binary mixtures. x_1 is the molar fraction of the first compound in the mixture.

Mixture	x_1 **range**	**SD** (average on the composition)	
		A_i **from Eq. 3.32**	A_i **from Eq. 3.33**
water + *n*-propanol	0.5-0.98	6.99	4.49
water + ethylene glycol	0.80-0.88	6.75	0.98
water + 1,2-propanediol	0.32-0.98	11.37	4.85
ethyl acetate + benzene	0.09-0.89	0.46	0.26
acetone + benzene	0.12-0.92	1.05	0.54
SD		5.32	2.22

The use of the Paquette equation for the calculation of A_i gives better results.
The mixtures containing water show large errors (mainly with Eq. 3.32), but also the authors (Suarez *et al.*, 1989) refer to very high errors in case of aqueous solutions exhibiting differences between the pure compounds surface tension of more than 20 mN/m.

3.5.2 Viscosity

The group contribution based model developed by Cao *et al.* (1993) can be used to predict the dynamic viscosity as:

$$ln(\eta V) = \sum_{i}^{NC} \varphi_i \, ln(\eta_i V_i) + 2\sum_{i}^{NC} \varphi_i \, ln\left(\frac{x_i}{\varphi_i}\right) - \sum_{i}^{NC}\left(\frac{q_i np_i \varphi_i}{r_i}\right)\sum_{j}^{NC} \theta_{ji} \, ln(\tau_{ji}) \qquad (3.34)$$

η (mPa·s) is the mixture viscosity; np_i is the proportionality constant of compound i; V is (cm^3/mol) the mixture volume, calculated as:

$$V = \sum_{i}^{NC} x_i V_i \qquad (3.35)$$

In Eq. 3.35 it is assumed that the excess value of the free volume is zero; η_i (mPa·s) and V_i (cm^3/mol) are the pure compound viscosity and molar volume. The parameters r_i and q_i are defined as:

$$r_i = \sum_{i}^{NC} \upsilon_{k,i} R_k \qquad (3.36)$$

$$q_i = \sum_{k}^{NC} \upsilon_{k,i} Q_k \qquad (3.37)$$

R_k, Q_k are group parameters obtained from the Van der Waals group volume (V_{wn}) and surface areas (A_{wn}) (Bondi, 1968); τ_{ij} is calculated from the group interaction parameter a_{mn}:

$$\tau_{ij} = exp\left(-\frac{a_{mn}}{T}\right) \tag{3.38}$$

The volume fraction φ_i and the parameter θ_{ij} are calculated as follows:

$$\varphi_i = \frac{x_i r_i}{\sum_j^{NC} x_j r_j} \tag{3.39}$$

$$\theta_{ji} = \frac{\theta_j \tau_{ji}}{\sum_L^{NC} \theta_l \tau_{li}} \tag{3.40}$$

The surface area fraction θ_j is defined as:

$$\theta_j = \frac{x_j q_j}{\sum_I^{NC} x_i q_i} \tag{3.41}$$

The pure compound viscosities η_i can be calculated with the M&G GC$^+$ model discussed in §3.2.
Table 3.15 shows the results for the mixture viscosity prediction (300 K). In the last row the standard deviation is reported. Good predictions for the mixture viscosity can be achieved.

Table 3.15. Squared residuals (R^2) on the mixture viscosity predictions for some binary mixtures.

Mixture	x_1	R^2
		Cao *et al.* (1993)
acetic acid + ethyl acetate	0.40	0.00
ethanol + methyl acetate	0.70	0.11
cyclopentane + 1-propanol	0.30	0.13
2-propanol + benzyl alcohol	0.80	0.08
methyl tert-butyl ether + benzene	0.66	0.00
methyl cyclopentane + 2-pentanone	0.30	0.05
anisol + toluene	0.21	0.00
tetrabromoethane + 1-octanol	0.70	0.47
cyclohexane + n-heptane	0.65	0.00
acetone + ethanol	0.34	0.17
SD		0.32

3.6 The GC-Flory EoS

The GC-Flory Equation of State (EoS) was developed by Holten-Andersen *et al.* (1987), for the prediction of VLE for polymer-solvent systems. Chen *et al.* (1990) proposed a simplified version, while Bogdanic and Fredenslund (1994) revised the model parameters. Saraiva *et al.* (1995) tested the model for the prediction of LLE.
The GC-Flory EoS is a predictive model since the needed model parameters are estimated through a GC-based model.
In the GC-Flory EoS, the activity coefficients γ_i are given by combinatorial (γ_i^C), free-volume (γ_i^{FV}) and attractive (γ_i^R) contributions:

$$ln\,\gamma_i = ln\,\gamma_i^C + ln\,\gamma_i^{FV} + ln\,\gamma_i^R \tag{3.42}$$

The three contributions are calculated as follows:

$$ln\,\gamma_i^C = ln\frac{\varphi_i}{x_i} + 1 - \frac{\varphi_i}{x_i} \tag{3.43}$$

$$ln\,\gamma_i^{FV} = 3\left(1 + C_i^{DOF}\right) ln\left(\frac{\tilde{V}_i^{1/3} - 1}{\tilde{V}^{1/3} - 1}\right) - C_i^{DOF}\; ln\frac{\tilde{V}_i}{\tilde{V}} \tag{3.44}$$

$$ln\,\gamma_i^R = \frac{1}{2} zq_i \left\{ \frac{1}{RT}\left[\varepsilon_{ii}\left(\tilde{V}\right) - \varepsilon_{ii}\left(\tilde{V}_i\right)\right] + 1 - ln\sum_j^{NC} \theta_j\, exp\left(-\Delta\varepsilon_{ji} / RT\right) + \right.$$
$$\left. - \frac{\sum_j^{NC} \theta_j\, exp\left(-\Delta\varepsilon_{ji} / RT\right)\Delta\varepsilon_{ji}}{\sum_k^{NC} \theta_k\, exp\left(-\Delta\varepsilon_{ki} / RT\right)} \right\} \tag{3.45}$$

$\tilde{V}_i$ is the reduced volume (ratio between the molar volume V_i and hard core volume V_i^*) for compound i; $\tilde{V}$ is the reduced volume of the mixture; z is the coordination number. The volume fraction φ_i is calculated as:

$$\varphi_i = \frac{x_i V_i^*}{\sum_j^{NC} x_j V_j^*} \tag{3.46}$$

The surface area fraction θ_i is calculated as in Eq. 3.41. The surface area q_i is calculated as in Eq. 3.37. The energy interaction parameters ε_{ji} and $\Delta\varepsilon_{ji}$ are given by:

$$\varepsilon_{ji} = \frac{\varepsilon_{ji}^0}{\tilde{v}} \tag{3.47}$$

$$\Delta\varepsilon_{ji} = \varepsilon_{ji} - \varepsilon_{ii} \tag{3.48}$$

The molecular hard core volume V_i^* and the interaction energy ε^0_{ji} are calculated through group contribution expressions:

$$V_i^* = 1.448 \cdot 15.17 \sum_n^{NG} \upsilon_{n,i} R_n \tag{3.49}$$

$$\varepsilon^0_{ji} = \sum_m^{NG} \theta_{m,i} \sum_N^{NG} \theta_{n,i} \varepsilon_{nm} \tag{3.50}$$

In the above expressions, *i* and *j* refer to compounds, *n* and *m* to groups.
ε_{nm} is defined as:

$$\varepsilon_{nm} = -\left[\varepsilon_{mm} - \varepsilon_{nn}\right]^{1/2} + \Delta\varepsilon_{nm} \tag{3.51}$$

C_i^{DOF} is the molecular external degrees of freedom parameter (temperature-dependent) defined as:

$$C_i^{DOF} = \sum_n^{NG} \left[C_{T0,n} + C_{T,n} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) + \sum_n^{NG} \frac{R_n}{\sum_m^{NG} R_m} C^0_n \right] \tag{3.52}$$

$C_{T0,n}$, $C_{T,n}$ and C^0_n are group parameters attributed to group *n* in molecule *i*, and T_{ref} is a reference temperature taken equal to 298.15 K.

3.7 The FV-UNIQUAC model

The FV-UNIQUAC model was developed by Bogdanic and Vidal (2000). The parameter table was than extended by Bogdanic (2001). The model is derived from the entropic-FV model (Elbro *et al.*, 1990; Kontogeorgis *et al.*, 1993). It associates the non ideality of a polymer solvent mixture with polymer segment-solvent interaction parameters. The polymer segments are the monomer repeating units of a polymer or copolymer. A solvent is a single segment. The activity coefficient is given by an entropic-free volume ($\gamma_i^{entr\text{-}FV}$) and a residual ($\gamma_i^R$) contribution:

$$ln\,\gamma_i = ln\,\gamma_i^{entr-FV} + ln\,\gamma_i^R \tag{3.53}$$

The entropic-free volume contribution is given by:

$$ln\,\gamma_i^{entr-FV} = ln\left(\frac{\varphi_i^{FV}}{x_i}\right) + 1 - \frac{\varphi_i^{FV}}{x_i} \tag{3.54}$$

The free volume (V_i^{FV}) is the calculated as:

$$V_i^{FV} = V_i - V_i^{**} \tag{3.55}$$

with the volume V_i^{**} defined as:

$$V_i^{**} = \frac{V_i^*}{1.448} = 15.17 \cdot \sum_n^{NG} \upsilon_{n,i} R_n \tag{3.56}$$

where V_i^* is the hard core volume of compound i (Eq. 3.49). The free volume segment fraction corresponds to:

$$\varphi_i^{FV} = \left(\frac{x_i V_i^{FV}}{\sum_J^{NS} x_j V_j^{FV}} \right) \tag{3.57}$$

The residual term (γ_i^R) is the same as UNIQUAC (Abrams and Prausnitz, 1975), taking into account that the polymer is a repetition of segments, and that in a copolymer different segments constitute the polymeric chain:

$$ln\,\gamma_i^R = \sum_k^{NS} \upsilon_{k,i} \cdot \left(ln\,\Gamma_k - ln\,\Gamma_{k,i} \right) \tag{3.58}$$

$\upsilon_{k,i}$ is the number of segments k in the component i. For an homopolymer $\upsilon_{k,i}$ equals to the number of repeated units, while for a solvent $\upsilon_{k,i} = 1$. NS is the number of segments. For a copolymer, $\upsilon_{k,i}$ is calculated as follows:

$$\upsilon_{k,i} = \frac{Mw_i}{\sum_m^{NS} X_{m,i} Mw_m} X_{k,i} \tag{3.59}$$

Mw_i is the molar mass of the copolymer, and Mw_m is the molar mass of each segment constituting the copolymer, and $X_{k,i}$ (or $X_{m,i}$) is the molar fraction of each segment. X_k is calculated as:

$$X_k = \frac{\sum_i^{NC} x_i \upsilon_{k,i}}{\sum_j^{NC} \sum_m^{NS} x_j \upsilon_{m,j}} \tag{3.60}$$

The activity coefficient of the segment (Γ_k) is calculated as in the UNIQUAC model (Abrams and Prausnitz, 1975). For a binary mixture:

$$ln\Gamma_k = Q_k\left[1 - ln\left(\sum_m^{NS}\Theta_m\tau_{mk}\right) - \sum_m^{NS}\frac{\Theta_m\tau_{km}}{\sum_n^{NS}\Theta_n\tau_{nm}}\right] \tag{3.61}$$

Q_k represents the surface parameter of the segment k (Bondi, 1968). Θ_n represent the segment surface fraction:

$$\Theta_m = \frac{X_m Q_m}{\sum_m^{NS} X_n Q_n} \tag{3.62}$$

The values of τ_{nm} are derived from the segmental interaction parameter ΔU_{nm}, with the relation:

$$\tau_{nm} = exp\left(-\frac{\Delta U_{nm}}{RT}\right) \tag{3.63}$$

The segmental interaction parameters are assumed to have a linear temperature dependency:

$$\Delta U_{nm} = a_{nm} + b_{nm}\left(T - T_{ref}\right) \tag{3.64}$$

m and n are the segments and T_{ref} is a reference temperature, which can differ for each segment pair (in this work, usually 273.15 K).
The activity coefficient of the segment k in the pure component i ($\Gamma_{k,i}$, Eq. 3.58) is equal to one for a solvent or for an homopolymer, while for a copolymer $\Gamma_{k,i}$ is calculated employing Eq. 3.61 where the surface fractions are obtained from the mole fraction of the segment in the copolymer ($X_{k,i}$). $\Gamma_{k,i}$ takes into account the internal repulsive effects in copolymers.
The model requires, as input, the densities of the compounds involved, the van der Waals volumes, the molecular surface parameters and the segmental interaction parameters.

4

METHODS & TOOLS

This chapter focuses on the methods (algorithms) and tools (computer programs) developed in this work for the solution of problems involving design and verification of formulated products. The methods here represent algorithms for mixture classification and analysis, while the tools are their implementation as computer programs.
Databases of chemicals and knowledge base of products and their attributes, etc, were also developed as part of this work. They are additional tools employed for the solution of the problems considered.
Section §4.1 is dedicated to the algorithms, while section §4.2 to the databases and knowledge base.

4.1 Algorithms

The algorithms developed in this work are:

- The mixture classification algorithm: it classifies the mixtures in terms of hydrogen-bonding properties (§4.1.1);
- The mixture design algorithm (MIXD): it designs mixtures that match a set of constraints (§4.1.2);
- The stability test algorithm (STABILITY): it tests the stability of liquid mixtures (§4.1.3).

4.1.1 Mixture classification algorithm

This algorithm classifies the mixtures according to the hydrogen-bonding concept (H-B), which was explained in Chapter 3, §3.4. The flow-diagram of actions is shown in Fig. 4.1. The algorithm is simple: the classification associates each solvent mixture (given as input) with one of the possible six combinations: NF/NF, PNA/PNA; PAS/PAS, NF/PNA, NF/PAS or PNA/PAS. If the mixture is of the first two kinds (NF/NF, PNA/PNA), excess properties of mixing are negligible. For all the other mixture types, the excess properties are not negligible and linear mixing rules are not reliable for the prediction of mixture properties.

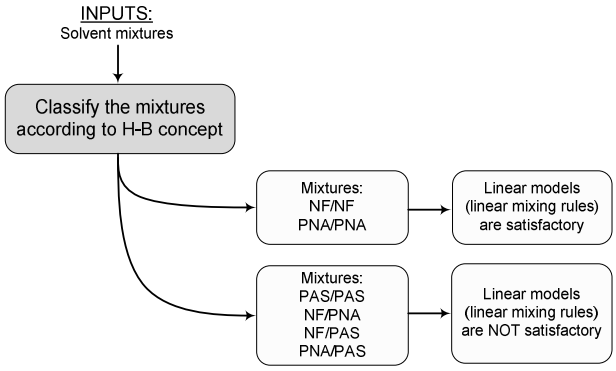


Figure 4.1. Flow-diagram of the mixture classification algorithm.

4.1.2 The MIXD algorithm

The mixture design problem being solved here is defined as: given a set of constraints on a set of target properties ($\underline{\zeta}$), determine the solvent mixtures that match the constraints.

The algorithm employs a decomposition based solution strategy where the number of feasible mixtures is continuosly decreased in subsequent levels (arranged in a hierarchy of calculations of increasing complexity). Four levels are involved, as shown in Fig. 4.2. The levels are related to the target properties for the desired mixture. The target properties are classified as linear (if linear models are used) and non-linear (if non-linear models are used). In addition, a rigorous stability test is performed to ensure the stability for the liquid mixture. The algorithm is described and highlighted for binary mixtures but can easily be extended to multicomponent mixtures.

The input information to the algorithm are: a solvent database where the necessary pure compound properties are stored, the property models for the description of mixture properties, the constraints on the target properties, the design temperature and information for the non-linear models employed.

The algorithm returns, as output, the mixtures matching the constraints, their composition, values of the target properties, cost, and information about the phase stability.

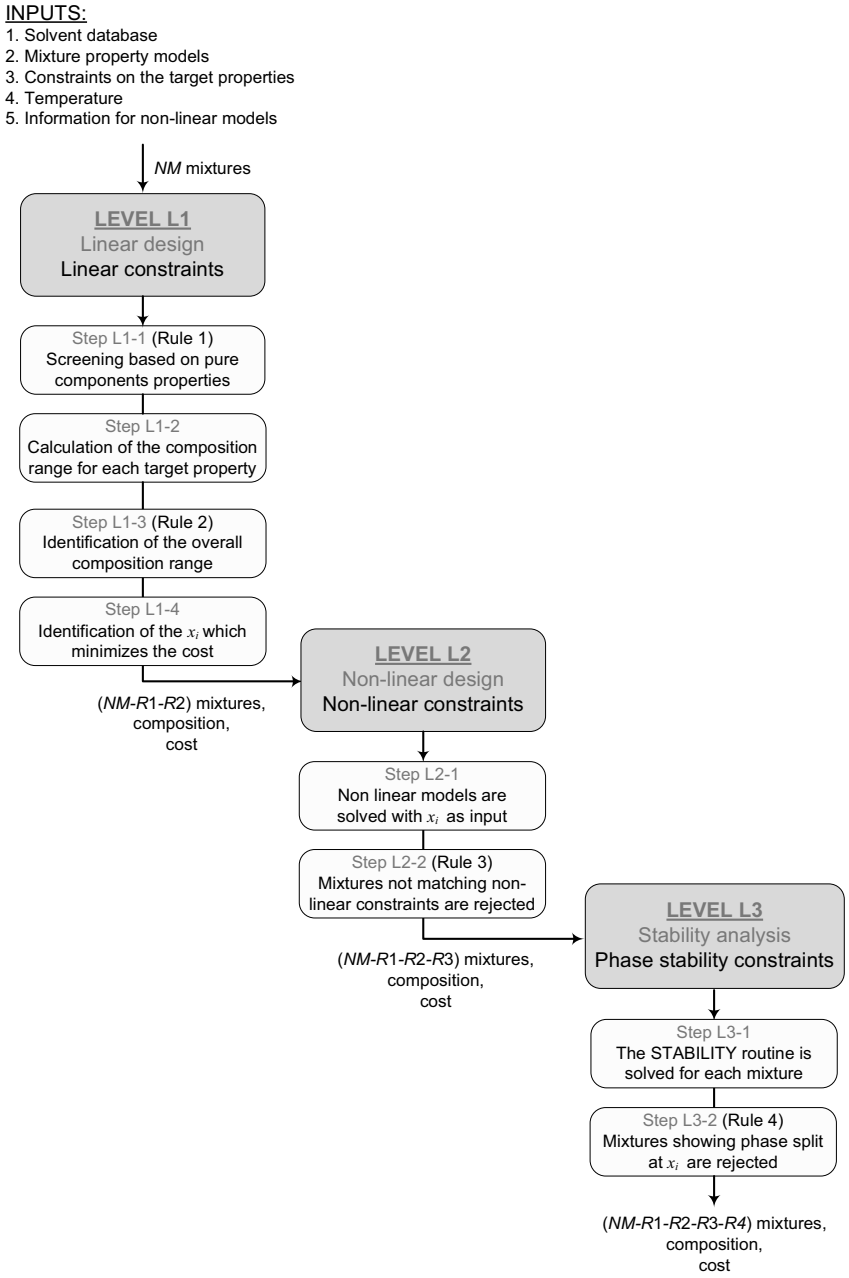


Figure 4.2. Flow-diagram of the MIXD algorithm used for the design of solvent mixtures.

The constraints on the generic target property ζ^k can be of the following types:

- Perfect match ($\zeta^{k,PM}$) constraints:

$$\zeta^k = \zeta^{k,PM} \tag{4.1}$$

When perfect match problems need to be solved, a slack of 0.5% is applied on the value $\zeta^{k,PM}$ to find the solutions as close as possible to the desired value. Hence, the equality constraints of Eq. (4.1) are transformed to inequality constraints of the type:

$$(0.995)\cdot\zeta^{k,PM} \le \zeta^k \le (1.005)\cdot\zeta^{k,PM} \tag{4.2}$$

- Upper ($\zeta^{k,UB}$) and lower ($\zeta^{k,LB}$) bound constraints:

$$\zeta^{k,LB} \le \zeta^k \le \zeta^{k,UB} \tag{4.3}$$

- Lower or upper bound constraints:

$$\zeta^k \ge \zeta^{k,LB} \tag{4.4}$$

$$\zeta^k \le \zeta^{k,UB} \tag{4.5}$$

In these cases, the upper bound in Eq. 4.4 is set to a very large positive value, and the lower bound in Eq 4.5 is set to a very large negative value.

It has to be underlined that, in a mixture design problem, there is a set of target properties (ζ) for each of the mixtures under consideration. Therefore, the vector ζ becomes a matrix, where each row represents a set of target property and each column a mixture. The matrix $\zeta(k\times m)$ is defined as shown in Eq. 4.6:

$$\zeta(k\times m) = \begin{vmatrix} \zeta^{1,1} & \zeta^{1,2} \\ \zeta^{2,1} & \zeta^{2,2} \\ & \cdots \end{vmatrix} \qquad \text{with } k = 1, NP;\ m = 1, NM \tag{4.6}$$

NP is the total number of target properties and *NM* the total number of mixtures (resulting from the combination of all the solvents in the database).
The element $\zeta^{1,2}$, for instance, is the mixture property number 1 for the mixture number 2.

4.1.2.1 Level L1

The screening of the mixtures starts at this level, in which linear constraints for the target properties are considered. Linear constraints are related to the properties described by linear models, which correspond to linear mixing rules (see Eq. 3.29 in Chapter 3).
For a binary mixture, Eq. 3.29 becomes:

$$\zeta^{k,m} = \sum_{i=1}^{NC} x_i \cdot \zeta_i^{k,m} = x_1 \cdot \zeta_1^{k,m} + (1 - x_1) \cdot \zeta_2^{k,m} \tag{4.7}$$

The subscripts 1 and 2 stand for compound 1 and compound 2 in the binary mixture.

Step L1.1

Rule 1. Reject a binary mixture if the pure component property values of both compounds in the mixture are either lower than the lower-bound values ζ_{LB} ($\zeta_1^{k,m} < \zeta^{k,LB}$ and $\zeta_2^{k,m} < \zeta^{k,LB}$), or greater than the upper-bound values ζ_{UB} ($\zeta_1^{k,m} > \zeta UB^k$ and $\zeta_2^{k,m} > \zeta UB^k$) of the target properties. ζ_{LB} and ζ_{UB} are vectors of dimension k.
Note. With $R1$ being the number of mixtures rejected after the application of rule 1, the total number of feasible mixture has been reduced to (MN-$R1$).

Step L1.2

Activity. Calculate the composition boundaries for each of the target properties considered in the design.
Note. For a general binary mixture m, the mixing process with respect to the general property k (ζ^k-axis), can be represented as shown in Fig. 4.3.

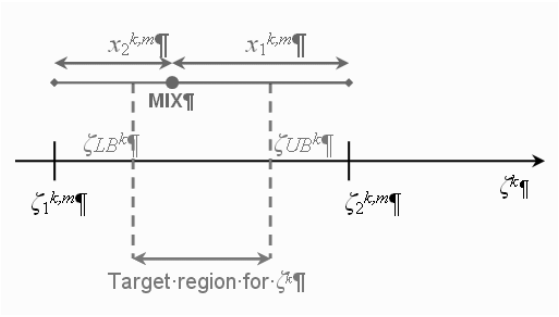


Figure 4.3. Representation of a binary mixing process with respect to the property ζ^k.

Solvent 1 and solvent 2 involved in the mixture have a position in the ζ^k-axis according to their pure component property values ($\zeta_1^{k,m}$ and $\zeta_2^{k,m}$); in Fig. 4.3 the case in which $\zeta_1^{k,m} < \zeta_2^{k,m}$ is shown. The target region is represented in red (dashed line), and corresponds to the segment between the lower bound (ζLB^k) and the upper bound (ζUB^k). The mixture property has to lie in the target region to match the design target: in Fig. 4.3 a possible solution is represented by the green dot. According to the lever rule, the composition of the mixture ($x_1^{k,m}$ since it is relative to the k-property and the m-mixture) can be calculated as follows:

$$x_1^{k,m} = \frac{\left(\zeta_2^{k,m} - \zeta^{k,m}\right)}{\left(\zeta_1^{k,m} - \zeta_2^{k,m}\right)} \tag{4.8}$$

x_1 is also a matrix: $x_1(k \times m)$.
When, instead of a single target value, bounds are given, Eq. 4.8 is substituted by Eqs. 4.9-4.10:

$$xLB_1^{k,m} = \frac{\left(\zeta_2^{k,m} - \zeta UB^k\right)}{\left(\zeta_1^{k,m} - \zeta_2^{k,m}\right)} \qquad \text{with } k = 1, NP;\ m = 1, NM\text{-}R1 \qquad (4.9)$$

$$xUB_1^{k,m} = \frac{\left(\zeta_2^{k,m} - \zeta LB^k\right)}{\left(\zeta_1^{k,m} - \zeta_2^{k,m}\right)} \qquad \text{with } k = 1, NP;\ m = 1, NM\text{-}R1 \qquad (4.10)$$

$xLB_1(k \times m)$ and $xUB_1(k \times m)$ are matrices containing the lower and upper bound values (respectively) of compositions calculated for all the target properties k and the mixture m:

$$xLB_1^{k,m} \le x_1^{k,m} \le xUB_1^{k,m} \qquad \text{with } k = 1, NP;\ m = 1, NM\text{-}R1 \qquad (4.11)$$

Eqs. 4.9-4.10 have to be solved for NP properties and (NM-$R1$) mixtures. From Fig. 4.3 it can be observed that the lowest value ($xLB_1{}^k$) that $x_1{}^k$ can assume in order to have the property ζ^k still matching the targets is where $\zeta^{k,m} = \zeta UB^k$. The highest value ($xUB_1{}^k$) that $x_1{}^{k,m}$ can assume in order to have the property ζ^k still matching the targets is where $\zeta^{k,m} = \zeta LB^k$.

Step L1.3

Activity. Identify the overall composition range for each mixture.
Note-a. The strictest condition for the composition has to be identified, for each mixture. That is, for each column of the matrix xLB_1 the maximum value has to be selected, while within each column of the matrix xUB_1 the minimum value has to be chosen, as shown in Eqs. 4.12-4.13:

$$xL_1^m = \max_{k=1}^{NP}\left(xLB_1^{k,m}\right) \qquad m = 1, NM\text{-}R1 \qquad (4.12)$$

$$xU_1^m = \min_{k=1}^{NP}\left(xUB_1^{k,m}\right) \qquad m = 1, NM\text{-}R1 \qquad (4.13)$$

$\underline{xL_1}$ and $\underline{xU_1}$ are vectors. The overall composition range (for all the mixtures) can be written using vectors:

$$\underline{xL_1} \le \underline{x_1} \le \underline{xU_1} \qquad (4.14)$$

Three different types of binary solvent mixtures can be classified according to the composition range evaluated at this step:

1. Type 1: binary mixtures for which the composition range of Eq. 4.14 is feasible, since $\underline{xL_1} \leq \underline{xU_1}$;
2. Type 2: binary mixtures for which the composition range of Eq. 4.14 is not feasible, since $\underline{xL_1} > \underline{xU_1}$;
3. Type 3: binary mixtures for which the composition range of Eq. 4.14 cannot be identified; this happens when the ranges defined by the different properties do not overlap. Consider, for example, that only two target properties are involved and that, for the mixture *m,* the first property constraint on $\zeta^{1,m}$ sets a composition range [0.2, 0.3], while the second property constraint on $\zeta^{2,m}$ sets a composition range [0.5, 0.7]: these two ranges do not overlap and no common composition range can be identified.

Rule 2. Reject mixtures of type 2 and 3.
Note-b. The composition range in which all the possible values of composition lead to a mixture of solvents that matches the design constraints has been identified at the end of step L1-3. With *R*2 being the number of mixtures rejected after the application of rule 2, the total number of feasible mixture has been reduced to (*MN*-*R*1-*R*2).

Step L1.4

Activity. Calculate the composition value that leads to the cheapest mixture, for all the mixtures that were not rejected with step L1-3.
Note. In the case of a binary mixture, the solution of this problem is straightforward. If, for mixture *m*, compound 1 is the cheapest between solvent 1 and solvent 2, the cheapest mixture is the one with the composition equal to xU_1^m, which is the maximum value of x_1^m (while still matching the constraints). If solvent 2 is the cheapest between solvent 1 and solvent 2, the composition that minimizes the cost is xL_1^m.

4.1.2.2 Level L2

At the end of level 1, (*MN*-*R*1-*R*2) binary mixtures matching the linear constraints, together with their composition and cost have been calculated. In this second level non-linear constraints are applied for further screening of the mixtures.

Step L2.1

Activity. Solve the non-linear models for each of the (*MN*-*R*1-*R*2) mixtures designed in level 1.
Note. The new mixture properties are now calculated employing the composition values calculated in the previous level.

Step L2.2

Rule 3. Reject the mixtures for which the property values do not match the non-linear constraints.
Note. With $R3$ being the number of mixtures rejected after the application of rule 3, the total number of feasible mixture has been reduced to (MN-$R1$-$R2$-$R3$).

4.1.2.3 Level L3

At the end of level 2, ($R1 + R2 + R3$) mixtures have been rejected. Only (MN-$R1$-$R2$-$R3$) mixtures match the non-linear as well as the linear constraints. Only these mixtures are now considered in level 3, for the stability test. At level L3, the STABILITY algorithm (§4.1.3) is employed. Phase split should not occur for the feasible binary mixtures.

Step L3.1

Activity. Apply the STABILITY algorithm on all the (MN-$R1$-$R2$-$R3$) feasible mixtures

Step L3.2

Rule 4. Reject the mixtures showing phase split at the design composition x_1.
With $R4$ being the number of mixtures rejected after the application of rule 4, the total number of feasible mixture has been reduced to (MN-$R1$-$R2$-$R3$-$R4$).

4.1.2.4 The MIXD program and application examples

A computer program based on the MIXD algorithm has been developed.
This software has been tested on three simple application examples. The same mixture design problem is considered in all the examples, to design at least one single phase binary solvent mixture combining the solvents of Table 4.1, which matches different constraints on the following properties: the toxicity parameter LC_{50}, the dynamic viscosity η, the molar volume V, and the solubility parameter δ. It has to be underlined that the examples just want to illustrate the algorithmfunctioning, but they do not exemplify the design of any particular product or solvent mixture.
The types of constraints applied on these properties are different in the three examples:

1. Example 1: perfect match constraints (Eq. 4.1) are applied on all the target properties;
2. Example 2: upper and lower bound constraints (Eq. 4.3) are applied on all the target properties;
3. Example 3: upper or lower bound constraints (Eq. 4.4-4.5) are applied on all the target properties.

Table 4.1. Small solvent database for the application examples of MIXD program.

nº	Name	UNIFAC-LLE decomposition	*Mw*	*C*	LC_{50}	*η*	*V*	*δ*
			kg/kmol	$/kmol	mol/l	cP	l/kmol	$MPa^{½}$
1	water	1 17	18.01	7.13	1.00E+00	0.89	18.00	47.84
2	ethanol	1 1 1 2 1 14	46.07	114.30	1.41E-01	1.08	57.30	26.52
3	2-propanol	1 16	60.09	85.51	6.76E-02	2.06	76.79	23.53
4	EGDME	2 1 2 28	90.12	484.30	2.24E-02	0.31	102.51	19.00
5	pentane	2 1 3 2	72.15	34.29	2.69E-03	0.24	116.34	14.00
6	1-Butanol	1 1 3 2 1 14	74.12	141.12	1.82E-02	2.64	92.05	22.54
7	2-Butanol	2 1 1 2 1 3 1 14	74.12	214.80	2.45E-02	3.17	92.05	23.35
8	1-Hexanol	1 1 5 2 1 14	102.17	198.22	2.29E-03	3.82	106.95	23.51
9	2-Hexanol	1 1 4 2 1 14	88.15	161.33	6.46E-03	3.09	92.95	24.22
10	1-Decanol	1 1 9 2 1 14	158.28	309.76	3.80E-05	8.94	161.91	22.09

Example 1

In this example the constraints on the target properties are:

$$LC_{50} = 0.13 \tag{4.15}$$

$$\eta = 1.21 \tag{4.16}$$

$$V = 60.00 \tag{4.17}$$

$$\delta = 26.10 \tag{4.18}$$

The results are shown in Table 4.2.

Table 4.2. Results of example 1: mixtures matching the targets and their property values.

nº	compound 1	compound 2	x_1	*C*	LC_{50}	*η*	*V*	*δ*	stability
				$/kmol	mol/l	cP	l/kmol	$MPa^{½}$	
1	ethanol	2-propanol	0.856	110.15	0.13	1.22	60.11	26.08	stable

Only one mixture matches the targets.

Example 2

In this example the equality constraints of example 1 are relaxed, and the following lower and upper bound constraints are applied:

$$0.14 \le LC_{50} \le 0.28 \tag{4.19}$$

$$0.70 \le \eta \le 1.21 \tag{4.20}$$

$$55.00 \le V \le 60.00 \tag{4.21}$$

$$25.50 \le \delta \le 26.50 \tag{4.22}$$

The results are shown in Table 4.3.

Table 4.3. Results of example 2: mixtures matching the targets and their property values.

n°	compound 1	compound 2	x_1	C	LC_{50}	η	V	δ	stability
				\$/kmol	mol/l	cP	l/kmol	$MPa^{½}$	
1	ethanol	2-propanol	0.984	113.85	0.14	1.09	57.61	26.47	stable
2	ethanol	EGDME	0.998	115.04	0.14	1.08	57.39	26.50	stable
3	ethanol	pentane	0.997	114.07	0.14	1.08	57.47	26.48	stable
4	ethanol	1-Butanol	0.996	114.40	0.14	1.08	57.43	26.50	stable
5	ethanol	2-Butanol	0.995	114.78	0.14	1.09	57.47	26.50	stable

Obviously, relaxing the constraints, more than one binary mixture matching the constraints is found.

Example 3

In this example the constraints of example 2 are relaxed: the constraint on the toxicity LC_{50}, the molar volume V and the solubility parameter δ now have an upper bound constraint of the type of Eq. 4.5; the constraint on the dynamic viscosity η becomes a lower bound constraint of the type of Eq. 4.4. The constraints for this example are:

$$LC_{50} \leq 0.28 \tag{4.23}$$

$$\eta \geq 0.70 \tag{4.24}$$

$$V \leq 60.00 \tag{4.25}$$

$$\delta \leq 26.50 \tag{4.26}$$

The results are shown in Table 4.4.

Table 4.4. Results of example 3: mixtures matching the targets and their property values.

n°	compound 1	compound 2	x_1	C	LC_{50}	η	V	δ	stability
				\$/kmol	mol/l	cP	l/kmol	$MPa^{½}$	
1	ethanol	2-propanol	0.862	110.31	0.13	1.21	60.00	26.10	stable
2	ethanol	EGDME	0.998	115.04	0.14	1.08	57.39	26.50	stable
3	ethanol	pentane	0.954	110.64	0.12	1.04	60.00	25.94	stable
4	ethanol	1-Butanol	0.996	114.40	0.14	1.08	57.43	26.50	stable
5	ethanol	2-Butanol	0.995	114.78	0.14	1.09	57.47	26.50	stable
6	ethanol	1-Hexanol	0.995	114.72	0.14	1.09	57.55	26.50	stable
7	ethanol	2-Hexanol	0.993	114.61	0.14	1.09	57.53	26.50	stable
8	ethanol	1-Decanol	0.997	114.96	0.14	1.10	57.66	26.50	stable

Once again, when relaxing the constraints, more mixtures become feasible solutions.

4.1.3 The STABILITY algorithm

The STABILITY algorithm checks the phase stability of a binary liquid mixture. The stability test is based on the trend of the Gibbs energy function of mixing ($\Delta G^{mix}/RT$)

and its first and second derivatives as a function of composition. The Gibbs energy of mixing is calculated as follows:

$$\frac{\Delta G^{mix}}{RT} = \frac{G^E}{RT} + \sum_{i=1}^{NC} x_i \cdot ln(x_i) \tag{4.27}$$

G^E is the excess Gibbs energy of mixing that is calculated from:

$$\frac{G^E}{RT} = \sum_{i=1}^{NC} x_i \cdot ln(\gamma_i) \tag{4.28}$$

Fig. 4.4 represents the four most common types of plots of $\Delta G^{mix}/RT$ as a function of x_i:

1. Mixtures of type *a* are completely immiscible in the composition range [0,1], and they can be recognized from the positive value of the function $\Delta G^{mix}/RT$ in the entire composition range [0,1];
2. Mixtures of type b_1 show a phase split in the composition range where the function $\Delta G^{mix}/RT$ is positive; the two liquid phase region corresponds to the region in which the $\Delta G^{mix}/RT$ is positive. The compositions of the two liquid phases are identified by the points in which the function $\Delta G^{mix}/RT$ is zero, at the extremities of the immiscibility gap;
3. Mixtures of type *c* are one phase in the entire composition range and they have negative values of the function $\Delta G^{mix}/RT$ and positive values of its second derivative in entire the composition range;
4. Mixtures of type b_2 are more complex: here, $\Delta G^{mix}/RT$ is negative and its second order derivative is negative between the compositions x_1^{δ} and x_1^{ε} (δ and ε are the wrong phases in equilibrium). These mixtures show a miscibility gap, but the compositions of the two liquid phases are not identified by the points in which the second derivative of $\Delta G^{mix}/RT$ changes its sign (x_1^{δ} and x_1^{ε}), since these points do not correspond to the composition at which the total Gibbs energy is at its global minimum. In order to identify the composition of the two liquid phases in equilibrium the tangent plane condition (Baker *et al.*, 1982) has to be employed.

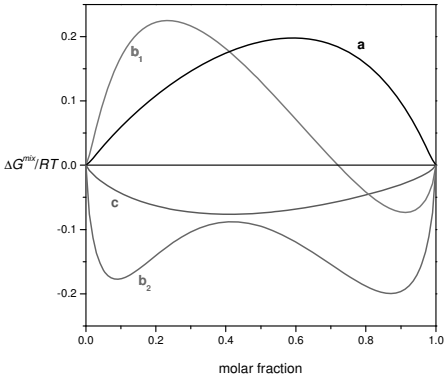


Figure 4.4. The most common shapes for the function $\Delta G^{mix}/RT$.

The fact that the second derivative of the function $\Delta G^{mix}/RT$ does not identify the 'real' immiscibility gap, while the tangent plane condition does, needs some explanations.

Fig. 4.5a shows the plot of $\Delta G^{mix}/RT$ and its first and second derivatives for the system ethanol + hexadecane (the system chosen for purpose of illustration). The second derivative becomes zero for $x_1^{\delta} = 0.62$ and $x_1^{\varepsilon} = 0.92$, but this range is not the 'real' two phase region, since the meta-stable regions (where the mixture is not stable) are not included in this range. Fig. 4.5b shows the plot $\Delta G^{mix}/RT$ and the meta-stable regions as well as the 'real' miscibility gap and the 'apparent' miscibility gap (the one identified analyzing the sign of the second derivative of the function $\Delta G^{mix}/RT$).

The points defining the 'real' immiscibility gap can be graphically illustrated as shown in Fig. 4.6, and they correspond to the tangent points found by drawing a tangent line ($y = at \cdot x_1 + bt$, with at and bt the slope and the intercept, respectively) to the $\Delta G^{mix}/RT$ surface, which has also to lie under the same $\Delta G^{mix}/RT$ surface. The mathematical expression for this condition is:

$$TPD = \frac{\Delta G^{mix}}{RT} - \left(at \cdot x_1 + bt\right) \geq 0 \qquad \forall x_1 \tag{4.29}$$

TPD is the tangent plane distance, which corresponds to the distance between the function $\Delta G^{mix}/RT$ and its tangent at every trial compositions.

The interpretation of this graphical solution from a thermodynamic point of view is the following: all the feed compositions z_1 in between the immiscibility gap $[x_1^{\alpha}, x_1^{\beta}]$ will split into two different phases of compositions x_1^{α} and x_1^{β} (α and β are the right phases in equilibrium) since the value of the Gibbs energy of mixing of the two phases is lower than the value at the feed composition z_1.

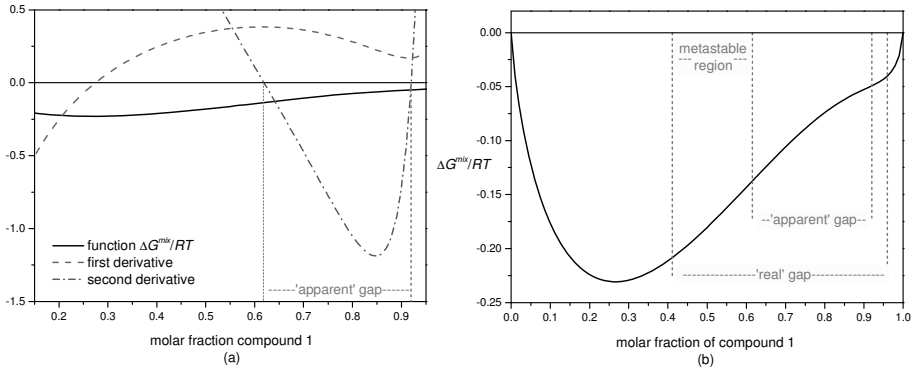


Figure 4.5. a) Function $\Delta G^{mix}/RT$ and its first and second derivative for the system ethanol(1) + hexadecane(2), at 298.15 K. b) Function $\Delta G^{mix}/RT$, 'real' immiscibility gap, 'apparent' immiscibility gap and meta-stable regions for the same system.

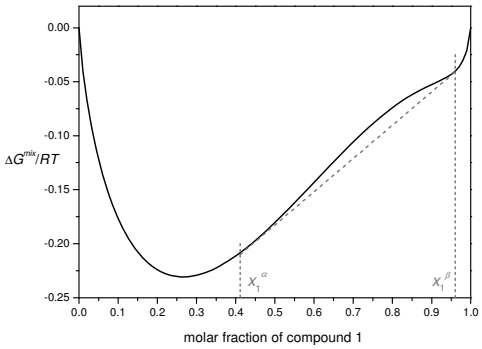


Figure 4.6. The real immiscibility gap is graphically identified by the tangent plane condition, for the system ethanol(1) + hexadecane(2), at 298.15 K.

Fig. 4.7 shows the work-flow of the STABILITY algorithm. The algorithm consists of three levels of screening. The UNIFAC model with the LLE group decomposition and contributions (Magnussen *et al.*, 1981) has been chosen to describe the binary chemical systems.

The algorithm needs the following information as input: the UNIFAC-LLE group decomposition of the chemicals involved in the mixtures under evaluation, and the temperature at which the stability test has to be performed.

The algorithm returns, as output, the stability information about the solvent mixtures (total miscibility, partial miscibility, total immiscibility), and, in case of partially miscible mixtures, the compositions of the two phases in equilibrium is also given.

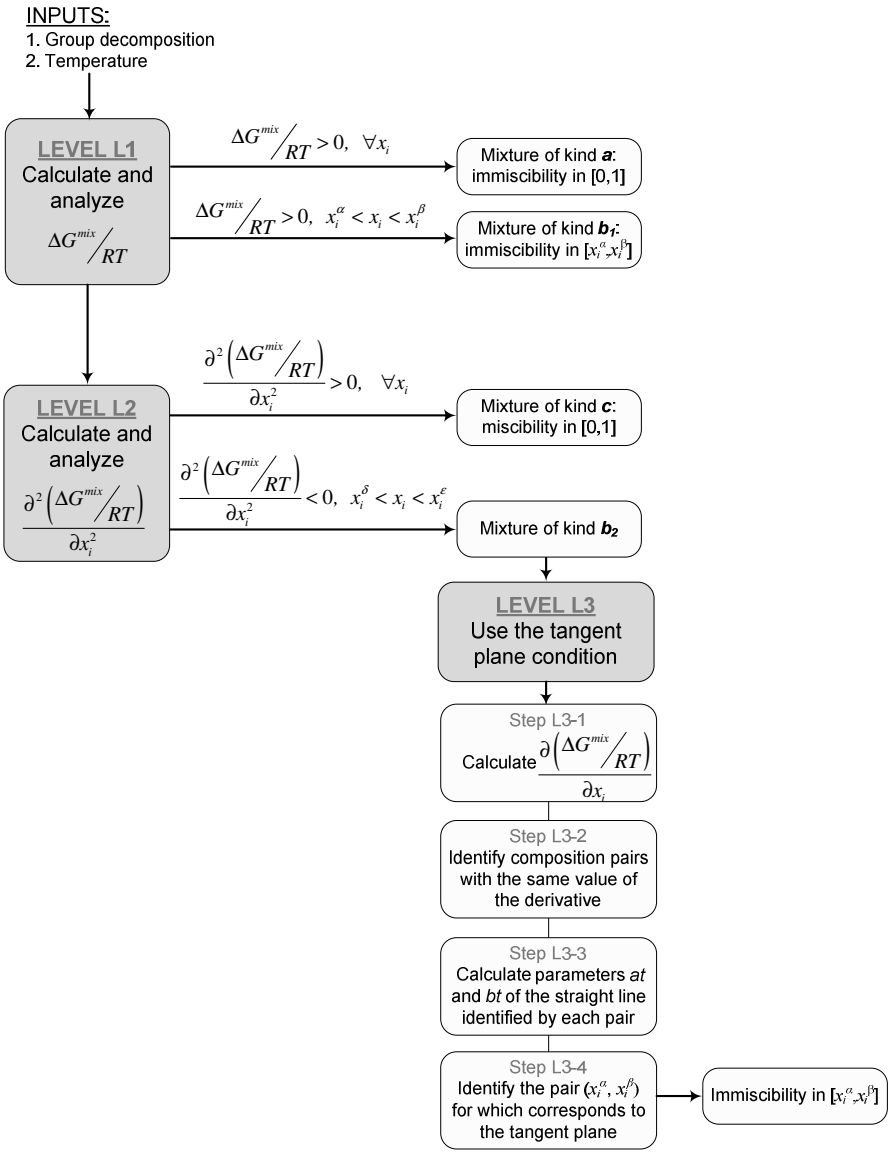


Figure 4.7. Work-flow diagram of the STABILITY algorithm.

4.1.3.1 Level L1

Activity. Calculate function $\Delta G^{mix}/RT$ and identify all the mixtures of type a and b_1. For mixtures of type b_1, calculate the immiscibility range $[x_1^{\alpha}, x_1^{\beta}]$ (it corresponds to the region in which the function $\Delta G^{mix}/RT$ is positive).

4.1.3.2 Level L2

Activity. Calculate the second derivative of the function $\Delta G^{mix}/RT$ and identify mixtures of type *c* (immiscible in all the composition range [0,1]) and b_2.

4.1.3.3 Level L3

For the mixtures of type b_2, employ the tangent plane condition to identify the immiscibility range $[x_1^{\alpha}, x_1^{\beta}]$.
Michelsen (1982) proposed an algorithm for the application of the tangent plane distance. This algorithm can be in principle used for any type of equilibrium (VLE, LLE,…), but in the case of liquid-liquid equilibrium the algorithm shows some drawbacks (Michelsen and Mollerup, 1981):

- In a liquid-liquid equilibrium, conditions are frequently 'near critical', and it is therefore often necessary to resort to 'second-order' procedures;
- The algorithm requires initial estimates of the *K*-factors (they correspond to the ratio of the composition of compound *i* in phase α and in phase β, that is, x_i^{α}/x_i^{β}), and no 'independent' *K*-factor estimate is available for LLE. The verification of stability by means of tangent plane analysis becomes much more complex and requires repeated calculations using different initial estimates.

In this work, the tangent plane condition is implemented following simple geometric rules (Step L3-1 to L3-4).

Step L3-1

Activity. Calculate the first derivative of the function $\Delta G^{mix}/RT$.

Step L3-2

Activity. Identify all the composition pairs with the same value of the derivative of the function $\Delta G^{mix}/RT$.

Step L3-3

Activity. Calculate the straight line determined by the *j* composition pair (from step L3-2) and repeat for all pairs. The equation for the straight line is:

$$y_1^j = at^j \cdot x_1^j + bt^j \tag{4.30}$$

Note. For every composition pair *j*, the values of the parameters *at* and *bt* (slope and intercept, respectively) are calculated solving the following simple system of equations:

$$\begin{cases} at^j = \dfrac{y_1^j - y_2^j}{x_1^j - x_2^j} \\ bt^j = y_1^j - at^j \cdot x_1^j \end{cases} \tag{4.31}$$

The tangent plane is one of the straight lines of Eq. 4.30.

Step L3-4

Activity. Identify the tangent plane between the *j*-straight lines of Eq. 4.30.
Rule 1. If Eq. 4.30 for the composition pair *j* satisfies the inequality of Eq. 4.32, it corresponds to the tangent plane.

$$TPD = \frac{\Delta G^{mix}}{RT} - \left(at^j \cdot x_1^j + bt^j \right) \geq 0 \quad \forall x_1 \tag{4.32}$$

Note-a. The meaning of Eq. 4.32 is: if the straight line *j* lies below the function $\Delta G^{mix}/RT$, it corresponds to the tangent plane, and the points of tangency are the compositions of the two phases in equilibrium, x_i^{α} and x_i^{β}.
Note-b. In correspondence of points of minima or maxima (or points where the function changes its convexity) of the function $\Delta G^{mix}/RT$, there are several composition pairs that satisfy Eq. 4.32, as shown in Fig. 4.8.

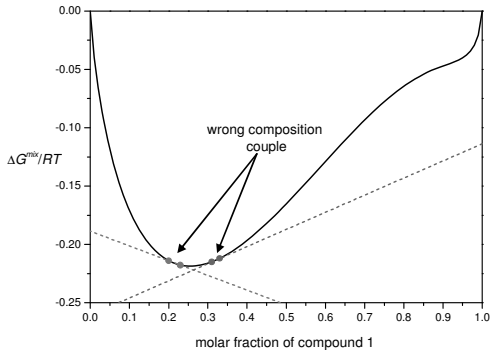


Figure 4.8. Wrong composition couples satisfying Eq. 4.32 for the system ethanol(1) + hexadecane(2). T = 298.15 K.

The composition pair, which shows the smaller difference of the derivatives of the function $\Delta G^{mix}/RT$, minimizing Eq. 4.33, is the one that identifies the tangent plane.

$$min \left| \left. \frac{\partial \left(\Delta G^{mix} / RT \right)}{\partial x_1} \right|_{x_1^{\alpha}} - \left. \frac{\partial \left(\Delta G^{mix} / RT \right)}{\partial x_1} \right|_{x_1^{\beta}} \right|^j \tag{4.33}$$

The concentrations of the two phases in equilibrium have been identified: x_i^{α} and x_i^{β} (immiscibility range [x_i^{α}, x_i^{β}]).

4.1.3.4 The STABILITY program and application examples

A computer program based on the STABILITY algorithm has been developed and tested for the prediction of the phase behaviour of liquid binary mixtures for which experimental data are available in the open literature.
Table 4.5 gives the results of the predictions from the STABILITY program and the known experimental values. The discrepancies between the predicted and the experimental values of the compositions of the two phases in equilibrium depend on the accuracy UNIFAC model (Magnussen *et al.*, 1981) used. Miscibility gaps are, however, correctly predicted for all the mixtures.

Table 4.5. Case studies for testing the performance of the STABILITY program. 'Predictions' are the results from the algorithm. In the last column the references for the experimental data are listed. Where the string 'miscible' appears, it means that no phase separation occurs.

nº	Systems	T (K)	predictions		experiments		References
			x_i^{α}	x_i^{β}	x_i^{α}	x_i^{β}	
1	water + pentane	293.20	0.002	0.999	0.000	1.000	Malagoni and Franco, 2007
2	ethanol + hexadecane	298.15	0.411	0.960	0.292	0.960	Hwang *et al.*, 2007a
3	ethanol + hexadecane	306.50	0.430	0.955	0.390	0.940	Browarzik, 2005
4	methanol + *n*-hexane	305.03	0.290	0.835	0.212	0.756	Browarzik, 2005
5	water + ethanol	303.15	miscible		miscible		Gramajo de Doz *et al.*, 2006
6	methanol + cyclohexane	301.8	0.198	0.841	0.138	0.809	Browarzik, 2005
7	heptane + anyline	323.00	0.083	0.856	0.176	0.878	Browarzik, 2007
8	water + isobutylacetate	283.15	0.001	0.785	0.001	0.945	Cháfer *et al.*, 2008
9	propylvinylether + water	323.15	0.023	0.774	0.019	0.826	Hwang *et al.*, 2007a

4.2 Databases and knowledge base

A library of databases has been created as part of this work, collecting all the ingredients employed for the case studies described in Chapter 6 and 7. A knowledge base library has also been developed, containing all the information needed when designing or verifying the products considered in this work. A detailed description of why and where databases and knowledge base are needed during the developemnt of a formulated product is explained in Chapter 5 (see Tables 5.1-5.2), where the methodology is highlighted in detailed.
The databases and knowledge base are here described, as some of the tools employed in the methodology developed in this work (Chapter 5).

4.2.1 Databases

The databases have been developed according to the needs set by the case studies taken into consideration for the validation of the proposed methodology (for the design and verification of formulated products). These databases are:

- AIs databases; one database has been created for each of the product activities of the formulations designed in Chapters 6-7. Therefore, the following databases have been developed:
  - Pigments database (Van der Walle *et al.*, 1999; Ansmann *et. al.*, 2001): it contains a list of pigments, and their properties. Pigments are used, for example, in paint formulations or in sunscreen products as physical UV blocker for sun radiations;
  - Insect repellents database (Badolo *et al.*, 2004; Qui *et al.*, 1997; Debboun, Frances and Strickman, 2006): it contains a list of chemicals that are well known to repel mosquitoes and that are employed in insect repellent products;
  - UV-A blockers database (Ansmann *et. al.*, 2001): it contains chemicals that can block the UV-A sun radiations and that are therefore used in sunscreen products;
  - UV-B blockers database (Ansmann *et. al.*, 2001): it contains chemicals that can block the UV-B sun radiations and that are therefore used in sunscreen products;
  - Antioxidants database (Ansmann et. al., 2001): it contains chemicals which can prevent the formation of free radicals, which work to break down the collagen and elastin in the skin that in turn causes the skin to wrinkle, sag, age. They are employed in cosmetic products such as skin creams and sunscreen products;
  - Polymers database (Van der Walle *et al.*, 1999; Varco and Williams, 1986; Shernov, 1991; Shah and Fernandez, 1994; Morawsky and Martino, 1997): this database contains a list of polymers and copolymers. These AIs are employed in several products, for instance in paint formulations and in hair sprays. In paint formulations, they are able to bind the insoluble particles of pigments and to provide for a protective coating of the surface. In hair spray, they provide for the holding power, the curl retention, the shine and luster.
- Solvents databases:
  - Water insoluble solvents commonly used in paint formulations (Klein *et al.*, 1992);
  - Water soluble solvents commonly used in paint formulations (Klein *et al.*, 1992);
  - Water insoluble alcohols; alcohols are usually employed in insect repellent lotions;
  - Water soluble alcohols;

- Esters; esters are usually employed in sunscreen lotions and they are well known to be water insoluble;
- Solvents usually employed in hair spray formulations;
- Water.

- Additives databases:
  - Aroma database (Arctander, 1969): aromas are frequently used in formulations (mainly cosmetic formulations or lotions to apply on the skin) to mask the unpleasant scent of some of the other compounds or to enhance the end-use product properties;
  - Preservatives database (Ansmann et al., 2001): they are chemicals which can prevent the microbial growth or undesirable product changes. They are usually employed in products to be applied on the body;
  - Wetting agents database (Shanti and Clifford, 1993): they are chemicals which lower the surface tension. They are usually employed in products which have to be spread on a surface, such as paint formulations;
  - Moisturizing agents database: they are mainly esters (oils) which can improve the moisture of the skin. They are usually employed in skin lotions and creams (sunscreen lotions, day creams, etc).

4.2.2 Knowledge base

The data and information collected for the knowledge base and used in this work have been retrieved from literature, patents, patented products, insights and common sense. The researcher (or scientist) using the tool can also add his/her own specific ideas to extend the knowledge base. The knowledge base offers the following options:

- The identification of the performance criteria/consumer needs/product attributes set ($\underline{\Psi}$) usually required for some specific formulated products (insect repellent lotion, sunscreen lotion, paint formulation and hair spray);
- The identification of the set of target properties ($\overline{\zeta}$) that affects the performance criteria set (ψ); this knowledge base has been presented in Chapter 3 (Table 3.1); this translation procedure is very critical in the design of consumer oriented chemicals based products, as it is not straightforward to identify which physicochemical properties affect a specific performance criterion;
- The setting of the constraint values on the target properties ($\overline{\zeta}$). Also here, patented products are taken as reference for the setting of the constraint values, but the researcher can also decide to improve the constraints values, in order to improve the existing product;

- The identification of the qualities to enhance/promote/correct with the addition of additives, etc, for specific formulated products. Dispersing agents are added to help the dispersion of solids, in case they are contained in the formulation. Solubilisers are added to promote the solubilisation of the AIs. Stabilizing agents are added to ensure the formulation stability. Wetting agents lower the surface tension of the whole formulation and so to promote spread-ability on surfaces. Preservatives are added to avoid the decomposition by microbial growth or by undesirable chemical changes. Aroma compounds are widely used to enhance the scent of the formulation. Moisturizing agents are added to improve the soothing effect;
- All the other type of information which could be useful in the design and verification of formulated products.

4.3 ICAS tools

Some of the calculations required in this work have been performed with the software ICAS (Nielsen *et al.*, 2001; ICAS Documentation, 2001). The Integrated Computer Aided System (ICAS) consists of a number of toolboxes that help to efficiently solve a wide range of problems: CAPEC DataBase, the Computer Aided Molecular Design tool (ProCAMD), the Property Prediction tool (PropPred), the Modelling Tool (MoT) and many more. ProPred is a very useful tool for the calculation of pure compound properties since it has a large number of property models and group contribution models (Marrero and Gani, 2001) that allow the user to get the needed properties for a very wide range of chemicals and polymers.
Almost all the pure compound property models listed in Table 3.3 (Chapter 3, §3.1) are available in ProPred. The pure compound property models (M&G GC^+ models) developed in this work for the prediction of surface tension and viscosity for liquids were also added to ProPred (Satyanarayana, 2009).
The ICAS utility provides a fast calculation option for a range of problems that usually needs to be solved in the design and verification of formulated products. The options currently available in ICAS Utility are: single stage flash calculations, vapour-liquid saturation point calculations, organic SLE phase diagrams, organic LLE phase diagrams, separation efficiency curves (driving force diagrams), aqueous electrolyte phase diagrams and electrolyte toolbox (for mixed electrolyte systems).

EQUATION CHAPTER 5 SECTION 1

5

METHODOLOGY & FRAMEWORK

When designing a new 'consumer oriented chemicals based product', the objective of the design (what consumers want) is usually known, while the ingredients and the composition of the product are not known. When verifying a 'consumer oriented chemicals based product', the identities of most of the chemicals are known, because a new alternative has been proposed and it is necessary to quickly evaluate the product properties. This scenario is often encountered in industry: experts suggest a list of ingredients that could be blended together to obtain a formulation, based on a combination of past knowledge, insight and their expertise. The product, however, may not be feasible due to various issues such as phase stability, or wrong product attributes. Also, the best blend may not have been identified. The methodology proposed in this work takes into consideration both the scenarios described above and it is presented in this chapter.

In §5.1 the methodology is described in detail. In §5.2 the framework for implementing the methodology in a computer-aided system, is introduced. More details about the framework can be found in Appendix F.

5.1 Methodology

In §5.1.1 an overview to the work-flow representing the methodology, which is based on the combination of modelling and experiments, is given. The computer-aided design stage is presented in §5.1.2 for the design scenario and in §5.1.3 for the verification scenario. The stages of experimental planning and experimental validation are treated in §5.1.4 and §5.1.5, respectively.

5.1.1 Overview

In this work, a methodology that integrates model-based design methods with experiments for formulated product design and verification is presented. The methodology is inspired by the integrated approach proposed by Ng, Gani and Dam-Johansen (2007). Three main stages are involved in this methodology:

1. STAGE 1 (S1): the model-based computer-aided stage. Here, models and computer-aided tools are employed to reduce the search space and to provide a list of potential candidates in the case of the design of a new product, or to evaluate a short-list of product candidates for the verification scenario. At the end of this stage, a base case product formula is proposed.
2. STAGE 2 (S2): the experimental planning stage. Here, a detailed plan for the experiments to be performed in the next stage is developed. The main objective is to determine which product attributes should be verified experimentally, how the experiments should be performed to measure them, which experimental set-up is needed, and so on.
3. STAGE 3 (S3): the experimental validation stage. Here the experimental trial-and-error approach is employed in order to test the base case formula proposed at the end of stage 1 and, if necessary, to amend it in order to identify the final product formula.

The main objective of the methodology is to quickly and efficiently generate a list of promising candidates for final testing (and selection) by experiments. In this way, rather than use the experiment-based trial-and-error approach from the start, the valuable experimental resources are reserved for final selection and testing while validated model-based computer-aided tools are used for identifying the promising feasible candidates.
In order to generate and screen thousands of design candidates, a robust and efficient model-based computer-aided method to reach a solution in a relatively short time is needed. The methodology employs the 'reverse design' technique (Chapter 2, §2.4) of defining a target and then finding alternatives that match the target in stage 1, and the 'forward design' (trial-and-error) in stage 3.
Fig. 5.1 shows the work-flow diagram of the integrated method.

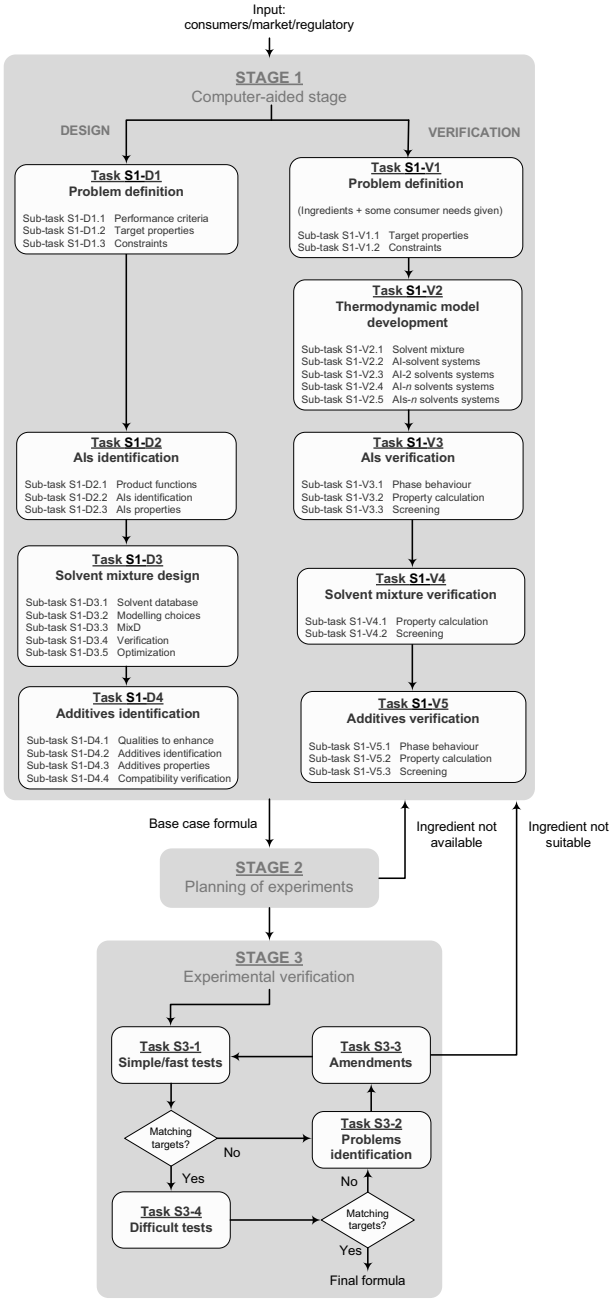


Figure 5.1. The integrated methodology for the design and verification of formulated products.

The following main characteristics can be observed in the work-flow diagram of Fig. 5.1:

- The model-based computer-aided stage (S1) does not involve iterations within the tasks and sub-tasks;
- The stage for planning of experiments (S2) may require iteration (in fact, an iterative loop connects stage 2 back to stage 1);
- The experimental stage (S3) may involve iterations within the stage itself and also with stage 1. In fact, an iterative loop connects stage 3 back to stage 1.

5.1.2 Computer-aided stage (S1): design (D)

Stage 1 for the design scenario performs model-based computer-aided design of chemicals based products. It consists of four main tasks, each of which involves several sub-tasks. Fig 5.2 shows a magnified part of Fig 5.1, only for the section of the work-flow related to the computer-aided stage, for the design scenario.

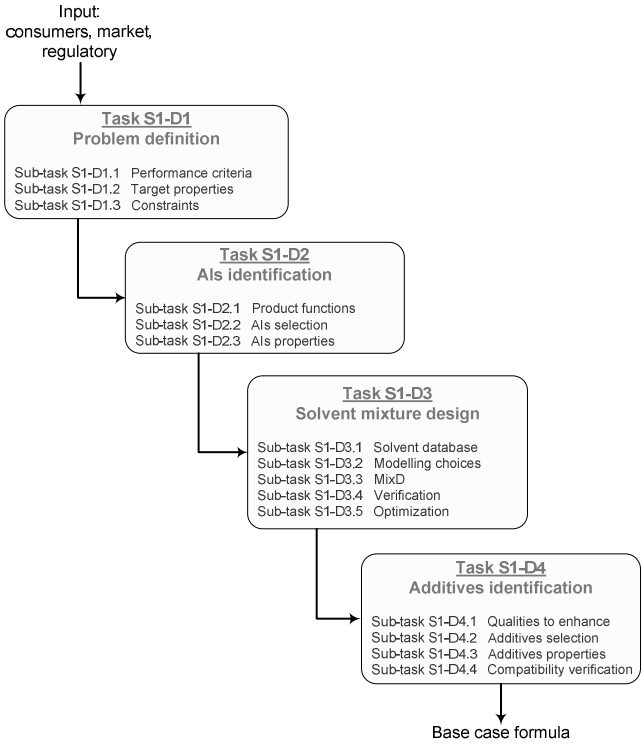


Figure 5.2. The enlargement Fig. 5.1 for the design scenario of the computer-aided stage.

5.1.2.1 Task S1-D1: problem definition

In this task, the performance criteria (consumer needs, product attributes) are identified, then they are translated into a set of physicochemical properties (target properties) and, finally, constraint values are set for the target properties.

Sub-task S1-D1.1: performance criteria

Activity. Identify a set of performance criteria ($\underline{\psi}$) for the product to be designed using the knowledge base (§4.2.2).

Note-a. The most important consumer need is the product main function: it is the main reason for which consumers buy the product. For instance, for an insect repellent, the main function is to repel mosquitoes. Consumer oriented chemicals based products can also have more than one main function. For instance, a sunscreen lotion has to protect the skin from ultraviolet radiation of type A and of type B, and it also has to prevent the skin aging; that is why in a sunscreen there are at least 3 active ingredients: a UV-A filter, a UV-B filter and an antioxidant.

Other performance criteria are related to the form of the product (liquid, solid, powder, spray,…), safety, drying time (if part of the product has to evaporate after application), cosmetic properties (if the product is to be applied on the body), and so on. For instance, if the product is a body lotion, consumers do not generally want a product that is sticky or with a bad scent.

Note-b. Some of the performance criteria will not be taken into consideration in this stage of the design (computer-aided), but only in the experimental planning and experimental validation. These properties are, for instance, the cosmetic and sensorial factors and the shelf life. In fact, these properties are difficult to describe through the use of target properties, while it is very easy and fast to verify them through experiments.

Sub-task S1-D1.2: target properties

Activity. Translate the performance criteria set ($\underline{\psi}$) into a set of target properties ($\underline{\zeta_i}$) using the knowledge base.

Note-a. The target properties $\underline{\zeta_i}$ are classified into three main types:

1. $\underline{\zeta_1}$: properties that determine the product main functions. These properties are related exclusively to the choice of the AIs.
2. $\underline{\zeta_2}$: properties that determine the product performance. They are the physicochemical properties. These properties are mainly related to the solvent mixture, and they can be only enhanced or corrected by the additives.
3. $\underline{\zeta_3}$: properties that determine the phase stability. They are the phase equilibria related properties. The stability of a liquid phase is controlled with the following conditions:

   3.1 Condition 1: the AI/AIs must be dissolved in the solvent mixture.

3.2 Condition 2: the solvent mixture must be a single liquid phase.
3.3 Condition 3: the additives must be soluble in the solvent mixture.

Note-b. It may not be possible to directly translate some performance criteria into chemical/physical properties even though they affect the design choices made later. For instance, the material compatibility with fabrics, metals and plastics is not translated into target properties, but it affects the selection of the ingredients.
Note-c. One performance criterion could be affected by more than one target property (see Chapter 3 where the translation process for the performance criteria has been given in Table 3.1).

Sub-task S1-D1.3: constraints

Activity. Set numerical constraints ($\underline{\zeta_i^{LB}}$, $\underline{\zeta_i^{UB}}$) on the set of target properties using the knowledge base (if necessary):

$$\underline{\zeta_i^{LB}} \le \underline{\zeta_i} \le \underline{\zeta_i^{UB}} \qquad \text{with } i = 1, 3 \qquad (5.1)$$

Assumption. Compounds having similar solubility parameters are miscible with each others (Hancock *et al.*, 1997).
Note. The assumption of Hancock *et al.* (1997) is employed to set (some of) the constraints for ζ_3, precisely the constraints for conditions 1 and 3 done in the previous sub-task (S1-D1.2):

$$\delta_{AI} - 3 \le \delta \le \delta_{AI} + 3 \qquad (5.2)$$

$$\delta_{AI} - 3 \le \delta_{add} \le \delta_{AI} + 3 \qquad (5.3)$$

δ is the solubility parameter of the solvent mixture, δ_{AI} is the solubility parameter of the AI (or average of the AIs solubility parameters, if more than one AI is present in the product), while δ_{add} is the solubility parameter of the additives. Eq. 5.2 corresponds to condition 1, while Eq. 5.3 corresponds to condition 3, as defined in sub-task S1-D1.2. To evaluate condition 2, the following constraints must be satisfied:

$$\frac{\Delta G^{mix}}{RT} < 0 \qquad (5.4)$$

$$TPD \ge 0 \qquad (5.5)$$

For more details on the Gibbs energy of mixing ($\Delta G^{mix}/RT$) and the tangent plane distance (TPD), see Chapter 4, §4.1.3).

5.1.2.2 Task S1-D2: AIs identification

The objective of this task is to retrieve from databases the chemicals that are able to satisfy the product functions ζ_1 defined at sub-task S1-D1.2. Then, the best performing chemicals in the list need to be selected. Next, the AI/AIs physicochemical properties that affect the next design tasks/sub-tasks must be retrieved from a database or calculated.

Sub-task S1-D2.1: product functions

Activity. For each of the desired main activities ζ_1 identified in sub-task S1-D1.2, generate a list of chemicals (AIs) by employing the AIs databases (Chapter 4, §4.2.1).
Assumption. The important hypothesis made here is that the AI/AIs are considered to be responsible only for the product main functions ζ_1, and that they do not strongly affect the performance criteria ζ_2.
This assumption is not valid if the AI/AIs are present with high concentrations or if the AI/AIs have a large contribution to one of the target properties. If this is the case, the selected AI may affect the other product qualities and this should be taken into consideration in the design. For instance, if the AI/AIs show high values of viscosity (honey-like), and viscosity is one of the target properties, the AI contribution to the final formula could be relevant. In this case, it is worth to define the constraint on the mixture viscosity taking into consideration the contribution of the AI/AIs.
For the kind of formulations considered in this work (formulation with a liquid form), the assumption made in this sub-task is very close to reality. In fact, in formulation with a liquid form, the solvent mixture is usually present in high concentrations. Hence the solvent mixture usually affects the product properties with the largest contribution, while the AI/AIs contribution is quite small, due to the low concentration.

Sub-task S1-D2.2: AIs selection

Activity. Select at least one chemical from each of the lists generated in sub-task S1-D2.1.
Note. The selection is done on the basis of one or more of the following criteria:

- Effectiveness: the most effective chemical is selected;
- Cost: the cheapest chemical is selected;
- Safety and health: the safest chemical is selected;
- Environment: the most environmentally friendly chemical is selected (for instance, VOCs emissions should be kept low);
- Material compatibility: the product ingredients should not dissolve plastics or damage fabrics, metals, and so on;

- Legislation: in some cases legislation needs to be examined since some chemicals are not anymore allowed for some applications or there are regulations about the concentration of AI in the product and so on;
- Others not listed above and/or combination of the above criteria.

Sub-task S1-D2.3: AIs properties

Activity. For each of the AIs selected in sub-task S1-D2.2, collect the necessary information/properties employing the databases (Chapter 4, §4.2.1) or calculate them by employing the property models (Chapter 3).
Assumption. Like dissolves like (Williamson, 1994).
Note. The necessary information/properties to be collected in this sub-task are:

- Information about the solubility in solvents (alcohol solubility, water solubility,…). These information are necessary when selecting the solvents for the solvent mixture design task (see sub-task S1-D3.1). For instance, if the AI/AIs are water soluble, only water soluble solvent need to be employed for solvent mixture design;
- The (Hildebrand/Hansen) solubility parameter. The solubility parameter is employed to fix the upper and lower bound values for the constraints of Eqs. 5.2-5.3, related to the product phase stability ($\underline{\zeta_3}$).

5.1.2.3 Task S1-D3: solvent mixture design

In this task, a set of candidate solvent mixtures that matches constraints on $\underline{\zeta_2}$ and $\underline{\zeta_3}$ are designed. If an additional verification of the mixtures is considered to be necessary, it could be performed here, too. At last, an optimization is carried on to identify the best performing solvent mixture according to the selected performance index (PI).

Sub-task S1-D3.1: solvent database

Activity. Select one or more solvent databases (consisting of solvents of a specific type and their properties) from the database library (Chapter 4), according to some of the performance criteria selected in sub-task S1-D1.1.

Assumption. Like dissolves like (Williamson, 1994).

Note. Since the solvent mixture should dissolve the AIs, solvents with solubility characteristics similar to those of the AIs (collected in sub-task S1-D2.3) have to be selected. For example:

- If the AI is well known to be alcohol soluble, the alcohols database is selected;
- If the AI is well known to be water soluble, only water soluble solvents are retrieved from the databases library. In this way, it is likely that the solvent

mixtures that will be designed in sub-task S1-D3.3 are able to dissolve the AI and are also one single liquid phase.

Sub-task S1-D3.2: modelling choices

Activity. Select the property models needed for predicting the mixture target properties from the model library (Chapter 3).
Note. As evident from Table 3.1 of Chapter 3, for properties such as viscosity and surface tension, the available property models are linear mixing rules or rigorous models. The later are based on the group contribution concept (viscosity: Cao *et al.*, 1993; surface tension: Suarez *et al.*, 1989). In the MIXD algorithm the models usually employed for the estimation of mixture viscosity and surface tension are linear mixing rules, since the rigorous models of Cao *et al.* (1993) and Suarez *et al.* (1989) are computationally expensive. The rigorous models are instead used in the verification sub-task (S1-D3.4), on the mixtures resulting from the MIXD algorithm.

Sub-task S1-D3.3: MIXD (solvent mixture design)

Activity. Apply the MIXD algorithm to design the solvent mixtures matching the constraints of Eq. 5.1 on ζ_2 and ζ_3 (sub-task S1-D1.3)
Assumption. As a consequence of the assumption made in sub-task S1-D2.1, all the constraints defined in sub-task S1-D1.3 on ζ_2 and ζ_3 are to be satisfied by the solvent mixture, with the exception of the constraint of Eq. 5.3 that is instead applied when selecting the additives (see task S1-D4). Note that the constraints on ζ_3 (Eqs. 5.4-5.5) are solved in the STABILITY algorithm which is part of MIXD.
Note. The mixtures matching the targets, their composition, cost, target properties values and stability information are the output of the MIXD algorithm.

Sub-task S1-D3.4: verification

Activity. Apply the mixture classification algorithm (Chapter 4, §4.1.1), verify the mixture target property values with rigorous models for the properties and mixtures that requires it and reject those mixtures that do not match the constraints on ζ_2 any more.

Sub-task S1-D3.5: optimization

Activity. Define a Performance Index (PI) and determine the optimal mixture by ordering all the feasible mixtures in terms of PI. For multiple PIs, determine the optimal solutions for each PI, than rank and weight the mixtures to identify the best 'trade-off' solution.
Note-a. At this stage a short-list of feasible candidates is available, therefore it is a simple task to order the feasible candidates according to a specific PI (PIs) to find the optimal selection.
Note-b. The actions performed in sub-tasks S1-D3.3 to S1-D3.5 are summarized in Fig. 5.3.

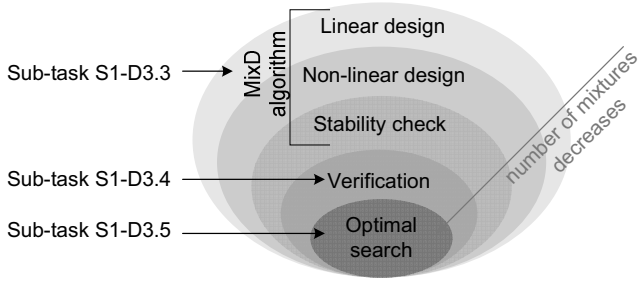


Figure 5.3. Hierarchy of activities performed in sub-tasks S1-D3.3 to S1-D3.5.

The hierarchy of activities in sub-tasks S1-D3.3 to S1-D3.5 reduces the number of feasible solvent mixtures. Every level employs a different set of constraints:

1. Linear design level (performed in the MIXD algorithm): linear property models (linear mixing rules) are solved and the mixtures that do not match the constraints on the target properties described by these models are rejected.
2. Non-linear design level (performed in the MIXD algorithm): non-linear property models are solved, using the results of the linear design level. The mixtures that do not match the constraints corresponding to the properties described with non-linear models are rejected.
3. Stability test level (performed in the STABILITY algorithm, which is part of MIXD): mixtures that show phase split at the composition of interest are rejected.
4. Verification level: rigorous models are employed, if necessary, to verify if mixture properties predicted with linear models (in level 1) are correct, and mixtures whose property values do not match the specified target values are rejected.
5. Optimization level: the optimal solvent mixture is selected according to any specific performance index (PI).

5.1.2.4 Task S1-D4: additives identification

The objective of this task is to add a final set of additives to the blend of AIs and solvents designed upto sub-task S1-D3.5.

Sub-task S1-D4.1: qualities to enhance

Activity. Identify the qualities to enhance or add to the formulation from the list available in the knowledge base.
Note. It could be necessary, for instance, to:

- Enhance the dispersion of solids;

- Promote the solubilisation of the AIs;
- Ensure the product stability;
- Enhance the spread-ability on surfaces;
- Avoid the microbial growth or by undesirable chemical changes;
- Enhance the sensorial factors and the cosmetic properties for products to apply on the body.

Sub-task S1-D4.2: additives selection

Activity. Identify a short-list of candidate chemicals for each of the qualities that have to be enhanced/added/promoted from the additives databases.

Sub-task S1-D4.3: additives properties

Activity. Retrieve from the additives databases the properties of the additives in the short-lists.
Assumption. Like dissolves like (Williamson, 1994).
Note. The necessary information/properties to be collected in this sub-task are:

- Information about the solubility in solvents (alcohol solubility, water solubility,…). For instance, if the additives are alcohol soluble, and the solvent mixture designed in task S1-D3 was a mixture of alcohols, the additives are more likely to be able to dissolve in the solvent mixture;
- The (Hildebrand/Hansen) solubility parameter. The solubility parameter is necessary for the next sub-task (S1-D4.4), where Eq. 5.3 is applied.

Sub-task S1-D4.4: compatibility verification

Activity. Apply the constraint of Eq. 5.3 and reject the additives that do not match this constraint.

Table 5.1 summarizes the content of each task/sub-task of the methodology of Fig. 5.2. It also shows the data flows between tasks/sub-tasks and the methods and tools to be employed.

Table 5.1. Data flows, models and tools used in the computer-aided stage for the design scenario (Fig. 5.2).

sub-task	input	action performed	models, databases, methods & tools	output
S1-D1.1	information about the product	understand user needs (performance criteria)	knowledge base[1]	list of performance criteria
S1-D1.2	list of performance criteria	translate the user needs into target properties	knowledge base[1]	list of target properties
S1-D1.3	list of target properties	set the constraints on the target properties	knowledge base[1]	list of constraints
S2-D2.1	activity of the product	find chemicals matching the activity set at S1-D1	database[1]	list of feasible AIs
S1-D2.2	list of feasible AIs, selection criteria	choose the most advantageous chemical for each activity	optimization methods	one AI for each product activity
S1-D2.3	AI/AIs identity	collect/calculate AI properties	databases[1], models[2], ICAS Utility[3]	solubility parameters, target properties, solubility info.
S1-D3.1	solubility/miscibility information, pure compounds properties	select a solvents database suitable for the AI	databases[1], models[2], STABILITY[3], ICAS Utility[3]	suitable property database
S1-D3.2	list of target properties	choose the models for the mixture target properties	linear/non-linear models for mixture properties[2]	mixture property models
S1-D3.3	databases, models, constraints, T	run the MIXD algorithm	MIXD[3]	feasible mixtures, properties, C, x_i,
S3-D3.4	feasible mixtures, properties, C, x_i, target properties to verify	perform product verification for critical target properties	H-B classification[3], rigorous mixture property models[2]	feasible mixtures, properties, C, x_i,
S1-D3.5	feasible mixtures, properties, C, x_i, PI	perform optimal search according to a PI	optimization methods	one optimal solvent mixture
S1-D4.1	AI/AIs + optimal solvent mixture	understand the qualities to add/enhance	knowledge base[1]	list of properties to enhance
S1-D4.2	list of properties to enhance	choose at least one chemical for each quality	databases[1]	list of possible additives
S1-D4.3	list of possible additives	collect/calculate the additives properties	databases[1], models[2]	solubility parameters, target properties, solubility info.
S1-D4.4	solubility parameters, target properties, solubility info.	verify compatibility with the solvent mixture	constraint Eq., ICAS Utility[3]	list of additives to add to the formulation

1) databases and knowledge base, Chapter 4
2) property models, Chapter 3
3) methods and tools, Chapter 4

5.1.3 Computer-aided stage (S1): verification (V)

The verification of a formulated product is a common scenario in industrial product development. Different situations could occur:

- A short-list of AIs is given and the objective of the verification is to identify the best performing AI (or blend of AIs);
- A short-list of solvents is given and the objective of the verification is to identify the best performing solvent/solvent mixture;
- A short-list of additives is given and the objective is to identify the best performing additive or blend of additives.
- Combination of the above situations.

It could also be necessary to calculate the product composition, or only the amount of one or few ingredients. For each of the above situations, different scenarios could also occur. This section gives the guidelines on how to tackle each different scenario.
Fig. 5.4 shows a magnified part of Fig 5.1, only for the section of the work-flow related to the computer-aided stage for the verification scenario

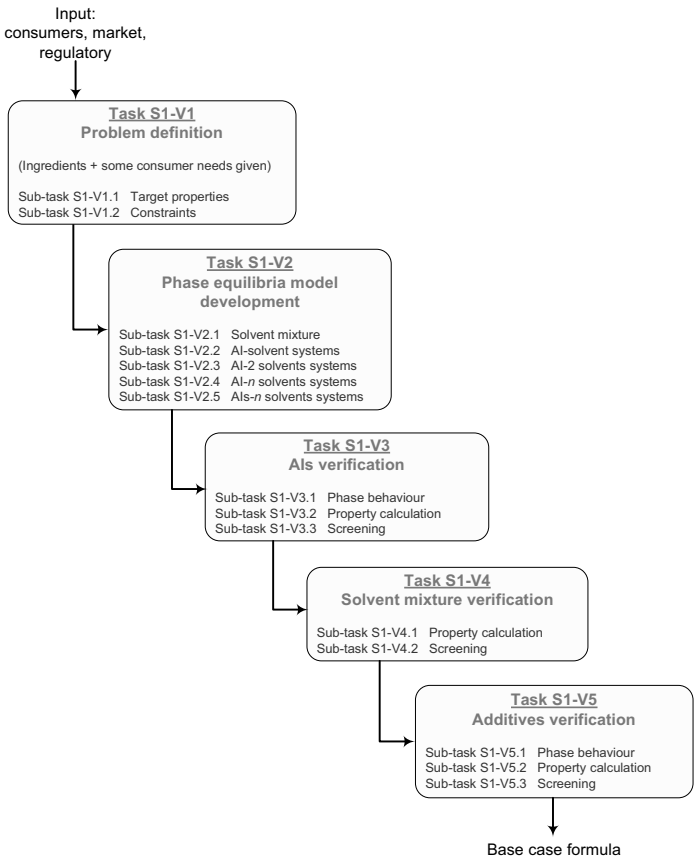


Figure 5.4. The enlargement Fig. 5.1 for the verification scenario of the computer-aided stage.

5.1.3.1 Task S1-V1: problem definition

The objective of this task is to define the verification scenario. How much of the formulated product is fixed (known)? How many and which target properties need to be verified?
Differently from the design of a formulated product, in the verification the performance criteria set ($\underline{\psi}$) is usually given.

Sub-task S1-V1.1: target properties

Activity. Translate the performance criteria set ($\underline{\psi}$) into a set of target properties ($\underline{\zeta_i}$).
Note-a. Also in the verification scenario, the target properties $\underline{\zeta_i}$ to be verified are classified into:

1. $\underline{\zeta_1}$: properties that determine the product main functions.
2. $\underline{\zeta_2}$: properties that determine the product performance.
3. $\underline{\zeta_3}$: properties that determine the phase stability.

Note-b. Unlike the design scenario, in the verification of formulated products it could happen that only some of the above target properties need to be verified. In fact, it could be necessary to focus just on the phase stability of the product, or just on the properties which determine the AI function, for instance.
Note-c. Another point of difference with the design scenario is related to the properties $\underline{\zeta_3}$. In the verification scenario, a rigorous phase stability property model for the description of the phase behaviour of AIs in solvents (and, if necessary, of additives in solvents) has to be developed/used.

Sub-task S1-V1.2: constraints

Activity. Set numerical constraints ($\underline{\zeta_i^{LB}}, \underline{\zeta_i^{UB}}$) on the set of target properties.

Note. The constraints on the properties $\underline{\zeta_3}$ (ΔG^{mix} and *TPD*) correspond to Eqs. 5.4-5.5. These constraints are now verified not only for the solvent mixture but also for the systems AIs + solvents, and/or AIs + solvents + additives.

5.1.3.2 Task S1-V2: phase equilibria model development

In this task, an appropriate phase equilibria model needs to be used and also developed if not available. Since formulations are complex chemical systems, models to predict properties of blends of AIs, solvents and additives are usually not readily applicable while experimental verification, although reliable, is time consuming and expensive. The objective, therefore, is to manage this complexity by breaking down the problem into a number of sub-problems where available data and/or models can be used.

Additives are not considered in this task since their concentrations are usually very low and therefore unlikely to influence the phase behaviour of the system. The additive effect on the phase behaviour is considered later, in task V5.
The problem decomposition (highlighted in Fig 5.5) starts with the analysis of the solvent mixture (sub-task S1-V2.1). Then the behaviour of one AI in each of the solvents is considered (sub-task S1-V2.2). One more solvent is added, and the ternary systems of AI + solvent binary mixture is considered (sub-task S1-V2.3). More solvents are then added and the AI phase behaviour in multicomponent solvent mixtures is considered (sub-task S1-V2.4). The case in which several AIs are present in the formula is also considered (sub-task S1-V2.5).

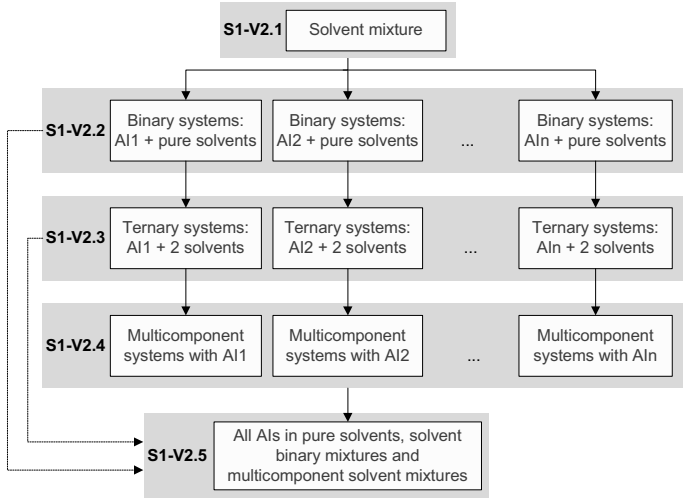


Figure 5.5. The sub-tasks (S1-V2.1 to S1-V2.5) of the phase equilibria task S1-V2.

The dashed arrows in Fig. 5.5 highlight the fact that after performing task S1-V2.2, the solubility of all the AIs in pure solvents can be simulated, or that after task S1-V2.3 the phase behaviour of all AIs in solvent binary mixtures can be calculated.

Sub-task S1-V2.1: solvent mixture

Activity. Model the phase behaviour of binary, ternary and multicomponent solvent mixtures.
Note. At first, a search of the databases and/or the open literature is performed to look for availability of experimental data covering the solvent systems of interest. Then, available property models (for instance UNIFAC, NRTL, or UNIQUAC) usually capable of predicting the liquid phase equilibria (LLE) for the systems under consideration are used to generate the phase diagrams for analysis. In this way, the solvents causing miscibility problems are identified.

Sub-task S1-V2.2: AI-solvent systems

Activity. Model the phase behaviour of the systems constituted of one AI and one solvent.

Note-a. A search of the databases is made to look for AI solubility information/data. Additionally, a search is also made in the open literature for more data, if necessary. At first, the applicability of available property models (from the library described in Chapter 3) based on group/segment contribution concepts to handle the chemical systems under consideration is tested with the available interaction parameters of the model. Note that the model employed in this sub-task could be different from the model/models employed in sub-task S1-V2.1. In fact, the AI/AIs involved in formulated products are usually big and complex chemical molecules, with multifunctional groups (APIs, polymers, pesticides,…). For systems involving this kind of molecules, models such as NRTL, UNIQUAC, UNIFAC do not guarantee success in the prediction of phase behaviour and physicochemical properties.

Note-b. After an appropriate model is selected, the problem of availability of parameters arises. If the existing parameters are not satisfactory, they are fine-tuned using experimental data. If the parameter values are not available, new parameters are regressed using experimental data. If experimental data are not available, pseudo-experimental data are generated with other models and then these data are used to regress the needed model parameters.

Note-c. After the model has been selected and the parameters regressed or fine-tuned, phase diagrams are calculated and the phase behaviour of sinlge AI in pure solvent is analyzed.

Sub-task S1-V2.3: AI-2 solvent systems

Activity. Model the phase behaviour of the systems composed of one AI and two solvent (binary solvent mixtures).

Note. Databases and/or literature are searched to find ternary solubility data involving pure AI and solvent binary mixtures. The model employed in this task is the same employed in the previous sub-task (S1-V2.2). If new binary or ternary interaction parameters are needed, the same procedure used in sub-task S1-V2.2 for the parameter fine-tuning or regression is employed. If no new parameters are required, the ternary experimental data can be employed to check/test the extrapolation of the phase behaviour in the ternary composition space (model validation).

Sub-task S1-V2.4: AI-n solvent systems

Activity. Model the phase behaviour of the systems composed of one AI and as many solvents as necessary (multicomponent solvent mixtures).

Note. At this point the behaviour of each AI in solvent mixtures with more than two solvents is taken into consideration. The model employed is the same used in subtasks S1-V2.2 and S1-V2.3. The procedure for model adaptation is the same as

above: literature search is performed and new parameters are fitted, or existing parameters are fine-tuned, if necessary.

Sub-task S1-V2.5: AIs-n solvent systems

Activity. Model the phase behaviour of the systems composed of all the AIs together and one solvent; then, AIs in solvent binary mixtures and multicomponent solvent mixtures.
Note. If several AIs are involved in the system, the procedure (from sub-task S1-V2.2 to sub-task V2.4) is repeated for each AI, as highlighted in Fig. 5.5.

5.1.3.3 Task S1-V3: AIs verification

With the phase equilibria model selected and verified (task S1-V2), this task focuses attention on the AIs.
According to the given information (see task S1-V1), different scenarios can occur, for example:

- Scenario 1: the AI/AIs are fixed (the identity of the AI/AIs is not under evaluation). The main function/functions of the AI/AIs are well known and they do not need to be verified.
- Scenario 2: the AI/AIs are fixed, but the main function/functions of the AI/AIs need to be verified.
- Scenario 3: at least one AI needs to be selected among the AIs short-list.
- Scenario 4: a blend of AIs needs to be identified. In this case, the amounts of AIs in the formula need also to be calculated.

Sub-task S1-V3.1: phase behaviour

Activity. Predict the phase behaviour of the AIs in solvents.
Note. It is necessary to predict the phase behaviour for all the scenarios being considered, to check for phase stability. For the model developed in task S1-V2, all the regressed interaction parameters are employed in this task.

Sub-task S1-V3.2: property calculation

Activity. According to the scenario being investigated, calculate the properties $\underline{\zeta_1}$.
Note-a. Further modelling effort for the prediction of the product main function/functions ($\underline{\zeta_1}$) may or may not be needed:

- Scenario 1: the product main function is known, therefore no extra modelling is required;
- Scenario 2-4: extra modelling is needed to calculate $\underline{\zeta_1}$. For instance, if the product is a hair spray, the main functions are: hair holding power and hair flexibility. The AI in a hair spray is usually a polymer (or copolymer). The hair

holding power is determined by the mechanical properties of the polymer. The hair flexibility is related to the flexibility of the polymeric chains. Hence, in this task it is necessary to calculate the polymer mechanical properties and chain flexibility.

Note-b. Since the screening has usually to be performed between few candidates, there is no combinatorial explosion of alternatives as in the design scenario (§5.1.2). Therefore, rigorous models (that take into account the excess properties of mixing) can be employed.
Note-c. If other calculations related to the AI/AIs are necessary, they are performed in this task. For instance, a sensitivity analysis may be needed, in some cases, to analyze the effect of the AI intrinsic parameters on the product functions or on the product stability (in the case of a polymer, molecular weight and molecular weight distribution strongly affect product functions and phase stability).

Sub-task S1-V3.3: screening

Activity. Reject those systems that do no match the constraints on ζ_3 and/or ζ_1.
Note. According to the possible scenarios, the constraints on ζ_1 may or may not be needed:

- Scenario 1: only the phase stability constraints (on ζ_3) are checked;
- Scenario 2-4: constraints on the phase stability (ζ_3) and on the product main functions (ζ_1) are checked.

5.1.3.4 Task S1-V4: solvent mixture verification

In this task, the solvents/solvent mixtures are considered. Different scenarios can occur also in this case:

- Scenario 5: the solvent or solvent mixture is fixed (identity of the ingredients and composition) and there is no need of verifying if the solvent/mixture satisfies the target property constraints;
- Scenario 6: the solvent/mixture is fixed but it is necessary to verify if the target property values match the constraints;
- Scenario 7: at least one pure solvent needs to be selected among the short-list of solvents;
- Scenario 8: a solvent mixture (combination of the solvents in a short-list) needs to be identified. Here, a composition that satisfies the constraints on the target properties needs to be calculated, if feasible.

Sub-task S1-V4.1: property calculation

Activity. If required by the given scenario, calculate the properties ζ_2.

Note-a. According to the scenario being investigated, further modelling effort for the prediction of the properties ζ_2 may or may not be needed:

- Scenario 5: no extra modelling is required;
- Scenario 6-8: extra modelling is required since the target property values need to be calculated.

Note-b. As for the previous task, rigorous models can be employed for the calculations of the properties ζ_2.

Sub-task S1-V4.2: screening

Actions. Reject those systems that do no match the constraints on ζ_2 and/or ζ_3.
Note. According to the possible scenarios, the constraints on ζ_2 and/or ζ_3 may or may not be needed:

- Scenario 5: no constraints need to be applied;
- Scenario 6-7: constraints on the properties ζ_2 need to be applied;
- Scenario 8: constraints on the properties ζ_2 and ζ_3 need to be applied. In fact, the composition of the blend AI (AIs) + solvent mixture is now known, and the LLE at this composition needs to be verified.

5.1.3.5 Task S1-V5: additives verification

In task V5, the influence of the additives on the formulation is analyzed.

Sub-task S1-V5.1: phase behaviour

Activity. Predict the phase behaviour of the additives in the systems AIs + solvents.
Note-a. Since the influence of the additives was not considered in the model development (task S1-V2), it now needs to be considered. The same model employed for the modelling of the AI + solvents systems are employed here. Phase diagrams need to be generated.
Note-b. Also in this case, different scenarios can occur:

- Scenario 9: the additive/additives are fixed, and there is no need of verifying the effect on the target property;
- Scenario 10: same as scenario 9, but it is required to quantify the effect of the additive/additives on the target property values;
- Scenario 11: one additive need to be selected from a short-list;
- Scenario 12: a blend of additives needs to be identified. In this case, the concentration of the additives in the formula need to be calculated.

Sub-task S1-V5.2: property calculation

Activity. If required by the scenario being investigated, calculate/recalculate the properties ζ_2.
Note. Further modelling effort for the prediction of the properties ζ_2 may or may not be needed:

- Sub-scenario 9: no extra modelling is required;
- Sub-scenario 10-12: extra modelling is required to quantify the effect of the additive/additives on the target property values.

Sub-task S1-V5.3: screening

Activity. Reject the systems that do not match the constraints on ζ_3 and/or ζ_2.
Note. According to the possible scenarios, the constraints on ζ_3 and/or ζ_2 may or may not be needed:

- Sub-scenario 9: only the phase stability constraints ζ_3 are checked;
- Sub-scenario 10-12: constraints on ζ_2 and ζ_3 are checked.

Table 5.2 summarizes the content of each task/sub-task of the methodology of Fig. 5.4. It also shows the data flows between tasks/sub-tasks and the methods and tools to be employed.

Table 5.2. Data flows, models and tools used in the computer-aided stage for the verification scenario (Fig. 5.4).

sub-task	input	action performed	models, databases, methods & tools	output
S1-V1.1	list of performance criteria	translate the user needs into target properties	knowledge base[1]	list of target properties
S1-V1.2	list of target properties	set the constraints on the target properties	knowledge base[1]	list of constraints
S2-V2.1	shortlist of ingredients, exp. data	modelling of solvent systems	models[2,3], STABILITY algorithm[3], ICAS Utility[3]	LLE solvent systems
S1-V2.2	shortlist of ingredients, exp. (pseudo) data	modelling of AI-pure solvent systems	models[2], ICAS Utility[2]	model parameters, LLE diagrams
S1-V2.3	experimental (pseudo) data	modelling of AI-solvent binary mixtures	models[2]	model parameters, LLE diagrams
S1-V2.4	experimental (pseudo) data	modelling of AI-multicomponent solvent mixtures	models[2]	model parameters, LLE diagrams
S1-V2.5	experimental (pseudo) data	modelling of AIs-pure solvents, multicomponent solvent mixtures	models[2]	model parameters
S1-V3.1	model parameters, LLE predictions from task S1-V2	generation of all necessary LLE phase diagrams	models[2]	LLE phase diagrams for the system AIs + solvents
S1-V3.2	identity of product main function	calculation of the necessary properties	models	values of main product functions
S1-V3.3	constraints on phase stability and product main functions	rejection of systems non matching	constraints Eqs.	feasible systems
S1-V4.1	identity of target properties	calculation of the target properties	models[2] (MIXD algorithm[3])	values of target properties
S1-V4.2	constraints on target properties (and phase stability)	rejection of systems non matching	constraints Eqs.	feasible systems
S1-V5.1	shortlist of ingredients	include the additives in the model previously developed	models[2]	LLE phase diagrams for the overall system
S1-V5.2	identity of target properties	calculation of the target properties	models[2]	values of target properties
S1-V5.3	constraints on target properties (and phase stability)	rejection of systems non matching	constraints Eqs.	feasible systems

1) databases and knowledge base, Chapter 4
2) models from the model library, Chapter 3
3) methods and tools, Chapter 4

5.1.4 Experimental planning stage (S2)

This stage constitutes the link between the computer-aided screening of alternatives and the experimental validation. Here, availability of chemicals is checked, experimental set-up is verified and decisions about the measurements and

experiments to be performed are achieved. If problems with the availability of chemicals are found, the opportunity of replacing the chemicals is investigated going back to stage 1 following the iterative loop shown in Fig. 5.1. If the experimental set up for a measurement or an experiment is not available, alternative solutions have to be found. At the end of this stage a list of experiments is produced.
The experimental planning divides the characterization experiments into three kinds.

1. The AI/AIs are tested in order to verify if they exhibit the needed functional properties for the product. For instance, if chemical A is chosen as the active ingredient in an insect repellent for tropical areas, it should be proven that it can repel the mosquitoes that are present in tropical areas. This validation work can be skipped if sufficient evidence could be obtained on the chemical function from literature or vendors.
2. The solvents and the solvent mixture are tested. The solvents should, above all, be able to dissolve the AI/AIs. Therefore liquid-liquid phase stability has to be tested. The physical/chemical properties of the solvents/solvent mixture that are considered critical for the product under consideration (as defined in the problem definition, task S1-D1 or S1-V1) should be measured, too. About the additives, their effect on the final product properties is expected to be negligible (since they are present in very small concentration), therefore they will not be tested individually but the effect on the overall formulation will be observed in the next level. Availability and price of these materials should also be considered.
3. Experimentation on the prototypes is performed. For the type of formulations considered in this work (with a liquid delivery system), the fabrication of the prototypes is simply mixing. The overall formula has to meet the *a priori* set targets. Some tests will include also the validation/measurements of properties that cannot be modelled, such as the sensorial factors and properties that have not been considered during the computer-aided design.

Experiments should be ordered according to complexity and time consumption. The simplest and least time consuming experiments should be the first ones in the list.

5.1.5 Experimental validation stage (S3)

In this stage of the methodology of Fig. 5.1, laboratory experiments are carried out in order to verify that the proposed formula is feasible. This stage constitutes of two iterative loops, the inner one is a slave loop since it is governed by the outer loop that can be defined as the master. In the inner loop the simple and non time consuming tests are carried on (task S3-1). If not all the tests are satisfactory, problems are identified (task S3-2) and amendments are carried on to fix the problems (task S3-3). Problems can be solved also going back to stage 1. In fact, if one AI out of many is not soluble in the solvent mixture, a new AI needs to be selected from the databases

of stage 1. The new AI solubility can be verified through the computer-aided tools of stage 1 and its properties can be calculated. After a suitable replacement for the AI has been found in the computer-aided stage, the experimental stage can be performed again.
The inner experimental loop is iterated until all the tests are satisfactory that is, when all the *a priori* defined constraints are satisfied. It may not be necessary to perform some of the tests once again. For instance, if the concentration of a chemical does not change from the n iteration to the (n+1) iteration, and the solvent mixture did not change either, it is useless to check the solubility of the AI in the solvent mixture once again.
When all the *a priori* defined constraints are satisfied, the inner loop of stage 3 is abandoned and task S3-4 is performed. Here all the difficult tests that require big employment of time and resources are carried on. If these tests are not satisfactory, the outer iteration loop has to be followed and the inner loop is entered again. Problems are identified and amendments are suggested (tasks S3-2 and S3-3). The simple tests have to be performed again since the amendments could affect the product properties considered in the inner loop (simple tests and measurements).
Only when the simple tests are satisfactory, the complicated tests of task S3-4 are performed again. This concatenation of loops is performed until the difficult tests are satisfactory (that implies that also all the simple tests are satisfactory).
When all targets are reached, a feasible formula has been identified. The next steps of product development can be now carried on, such as process synthesis and scale-up (not considered in this work).
It is evident that the experimental stage of product design is iterative. The computer-aided stage and the experimental stage are not independent between each other but integrated through the iterative loop going from task S3-3 to the last three tasks of stage 1, which involve ingredients selection. Some of the problems that can be found during the experimental stage can be solved just replacing one or more ingredients, and it becomes necessary to go back to the computer-aided stage where AIs, additives and solvents databases can be consulted, pure compound and mixture properties can be calculated, a new solvent mixture can be designed, solubility can be tested and so on.

5.2 Framework overview

Developing methods and tools for product design problems is almost as important as being able to integrate them efficiently in order to solve a wide range of product design problems, which involves different aspects and therefore different types of calculations. The development of a systematic framework for product and process design is one of the issues and needs discussed in Chapter 2, and it is one of the objectives of this thesis to address this need through the development of a computer-aided system that uses as its framework the methodology for design and verification

of liquid formulated products. The computer-aided stage of the methodology including all the related models, databases, methods and tools, has been implemented within an existing in-house software: the 'virtual Product-Process Design laboratory' (virtual PPD-lab) (Morales-Rodriguez, 2009), thereby extending its applicability range. The virtual PPD-lab is an Excel based software that provides a framework for the design/verification of chemical products and processes. The product developer can ask for the generation of a short-list of the most promising feasible product candidates (generation of alternatives). Alternatively, he/she can test his/her ideas on validated model-based computer-aided tools (that is, perform virtual experiments) rather than perform trials with experiments. The experimental resources, in this way, can be reserved for the final selection and validation step, thereby entering the market faster at a reduced cost.
The virtual PPD-lab offers the following advantages to the user:

- Problem decomposition into smaller sub-problems,
- Data management,
- Integration of methods and tools,
- Flexibility,
- Storage of past experience,
- User-friendliness.

In the virtual PPD-lab, product-process design problems are broken down into a hierarchical set of tasks/sub-tasks (problem decomposition) that are then integrated through efficient data management to reach a final design and/or evaluation of the overall system (integration). This allows the user to concentrate on the more important tasks of the design avoiding the manual and time consuming operations such as data collections, transfer of information, programming, data file generation, and so on. The software is also flexible since it allows the adaptation to different types of problems: in fact, the software hosts several tailor-made work-flows for the design of specific products, but, at the same time, the user can also introduce his/her own modifications, data, ideas, models, and so on. The choices/ideas/results are stored in output files that constitute the documentation for the developed case study (storage of past experience). The introduction of new data in the existing database or the creation of new databases is also contemplated. In addition, the software employs dialogue boxes to communicate with the user, and databases are Excel worksheet that can be directly modified and introduced by the user (user-friendliness).
The virtual PPD-lab hosts:

- Databases,
- Knowledge base,
- Property models,

- Process models,
- Methods (algorithms) and tools;
- Dedicated work-flows for specific product/process design/verification.

The virtual PPD-lab hosts a collection of databases that facilitate the developer in decision making such as selection of chemicals. Knowledge base gathered during the development of the case studies is also included. Models for property predictions are at the core of chemical product design, therefore the developer needs a reliable and flexible modelling framework so that necessary properties (pure component and mixture) for various types of chemical systems can be estimated and compared. The virtual PPD-lab provides this feature through its specially developed properties tool-box. It contains options for checking liquid miscibility, solid solubility, polymer solubility as well as bulk mixture properties such as density, viscosity, and evaporation rate. A library of predictive property models based on group-contribution (GC) and atom-connectivity (AC) methods is available. Process models are available in each work-flow. The architecture (framework) of the virtual PPD-lab allows the addition of new models, data and adaptation of existing models for future extension of application range of the virtual laboratory. In addition, the virtual PPD-lab can access various tools from ICAS (Nielsen *et al.*, 2001; ICAS Documentation, 2001), such as the CAPEC database, the CAMD tool (Computer Aided Molecular Design) and the pure component property prediction tool (ProPred). Tailor made work-flows for specific products and processes have been developed and incorporated into the virtual PPD-lab:

- Design of microcapsule for controlled release of Active Ingredients (AI) from a polymeric microcapsule (Morales-Rodriguez, 2009),
- Design of pesticide uptake in plants (Morales-Rodriguez, 2009),
- Design of direct methanol fuel cell (Morales-Rodriguez, 2009),
- Design and verification of formulated products (this work, §5.2.1).

Fig. 5.6 shows the software user interface, main menu screenshot (Morales-Rodriguez, 2009).

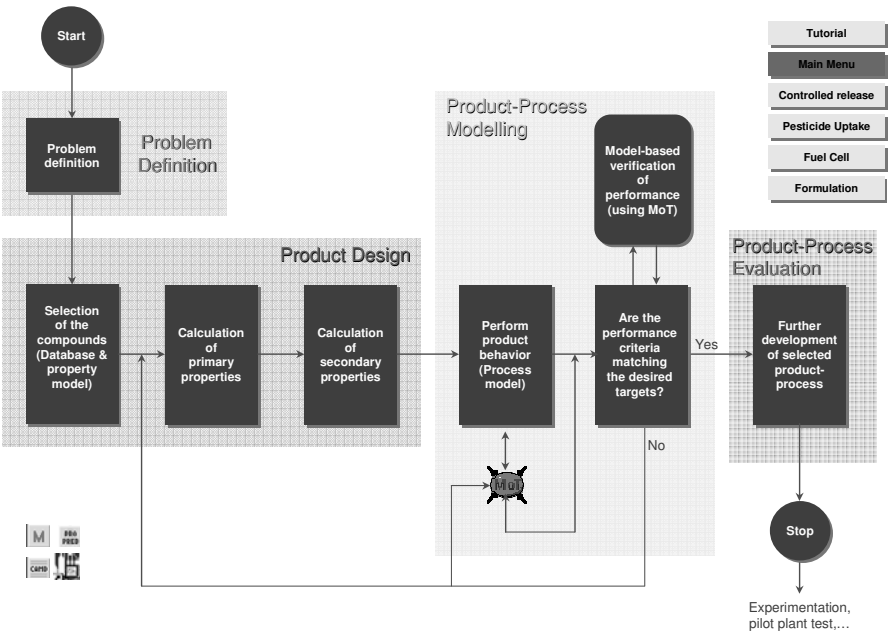


Figure 5.6. The user-interface of the 'virtual Product-Process Design laboratory'.

The software hosts sections for the overall problem definition, product design, product-process modelling and, at last, a section for the product-process evaluation (Morales-Rodriguez, 2009). In the software main menu (Fig. 5.6), general descriptions of the tasks performed in each section are shown (each box is a task), but the actions performed in each task change according to the product/process taken into consideration, which is chosen on the menu shown on the top right.

The problem definition section collects data and knowledge related to the product. In addition, new data and information about the specific case studies under development can be introduced: historical records, needs, assumptions, target settings, and so on. These information are needed in the next sections.

The product design section involves, in general, selection of the ingredients/materials for the product to be designed, calculations of primary properties (depending on the chemical structure, temperature and pressure) and functional properties (depending on the primary properties), as well as the calculation/simulation of the product properties/behaviours (such as phase behaviour, sensitivity analysis to some product parameters,…) needed in the next tasks. Product alternatives can be generated and evaluated. Databases, model libraries, toolboxes are employed in this step. If the necessary data/databases, models are not available, they can be generated by the user and included in the software libraries.

Once all the necessary information for evaluating the product performance have been retrieved, the product-process modelling section can be faced employing a modelling tool such as MoT (Nielsen *et al.*, 2001; ICAS Documentation, 2001). The simulation,

generation of process alternatives and verification of the *a priori* defined targets (set in the problem definition section) are performed here. If the product behaviour does not satisfy the target behaviour, the iterative loop shown in Fig. 5.6 can be followed: the process modelling can be performed a second time changing the model (in case the one previously employed was not adequate for the system under consideration) or changing some of the process parameters. When targets are matched, the designed product-process can be evaluated in the product-process evaluation section, before proceeding with the experimental part, pilot plant test and so on.

5.2.1 The formulation design feature

A new feature for the design and verification of formulated products with a liquid delivery system has been added to the virtual PPD-lab (shown in Fig. 5.7). The framework for this feature is based on the work-flow of the methodology described in §5.1.

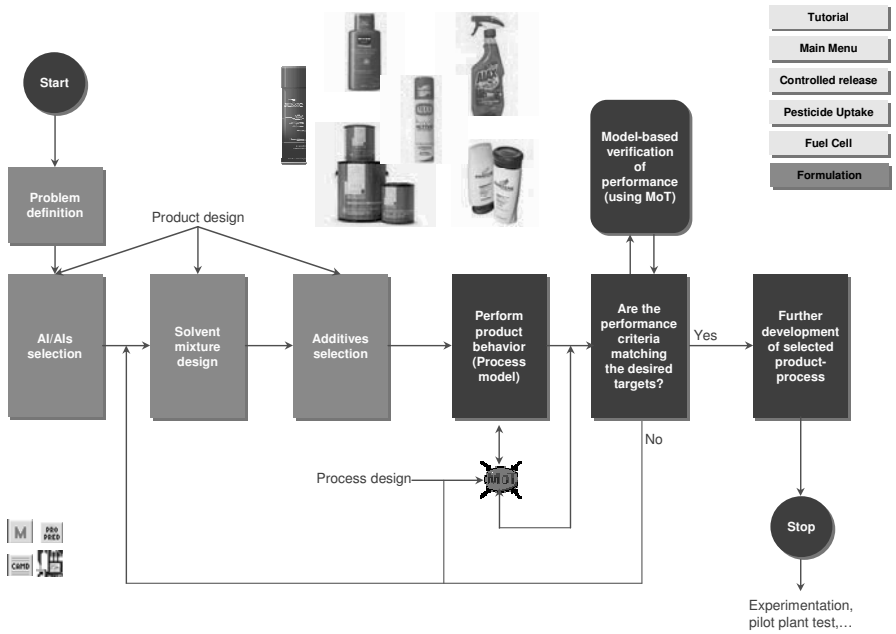


Figure 5.7. The new work-flow for the design and verification of formulated products in the virtual PPD-lab.

The work-flow involves the general sections of problem definition and product design (see Fig. 5.6). In the product design section, three tasks are included: the Active Ingredient selection, the solvent mixture design, and the additives selection. When the user selects the formulation work-flow from the top right menu, the above mentioned tasks are highlighted in light blue. The user has to follow the instructions asked in each task through dialogue boxes. Each task can be performed as a standalone task, i.

e, if the user only needs to search for additives, he/she can enter the additive selection task and search the databases. If the user wants to design a formulated product, all the four tasks of Fig. 5.7 need to be performed sequentially. The user can always amend his/her design and go back to the previous step, and change information, data, ingredients, constraints, and so on.
The work-flow for the design and verification of formulated products can be employed in several different 'modes':

- Design of formulations with a liquid delivery systems,
- Design of a solvent mixture,
- Database search,
- Calculation of pure compound and mixture properties,
- Verification of liquid-liquid equilibrium,
- Test of Active Ingredients solubility in pure solvent/solvent mixture,
- Test of additive solubility in pure solvent/solvent mixture.

The detailed description of the formulation work-flow structure is highlighted in Appendix F (user manual). An application example of the software (insect repellent case study) is given at the end of Chapter 6 (case study 4, §6.4).

EQUATION CHAPTER 6 SECTION 1

6

DESIGN CASE STUDIES

This chapter is dedicated to cases studies involving the design of formulated products. The application of the methodology is illustrated for the design scenario (see Fig. 5.1 in Chapter 5) involving the following three products:

- Insect repellent lotion (Conte *et al.*, 2009b), §6.1;
- Sunscreen lotion (Conte *et al.*, 2009b), §6.2;
- Paint formulation (Conte *et al.*, 2009a, 2010a), §6.3.

For the insect repellent lotion and the sunscreen lotion, the use of all three stages of the methodology (computer-aided design stage, planning of experiments stage, experimental validation stage) is shown. For the paint formulation, only stage 1 is highlighted as it was not possible to do the experimental validation.
An additional case study is presented in §6.4, where the application of the implemented work-flow in the 'virtual Product-Process Design laboratory' is illustrated. This case study involves the design of an alternative formulation for the insect repellent lotion (Conte *et al.*, 2009b, 2010a).

6.1 Case study 1: insect repellent lotion

The aim of this case study was to design an insect repellent lotion. Water should be one of the formulation ingredients, because of safety and cost concerns. Some well known insect repellents in the spray form are based on water-alcohol mixtures, like the well known product from Bayer, Autan®. The market for consumption is non tropical areas such as Europe.
An insect repellent lotion is usually constituted of the AI/AIs, with the function of repelling the mosquitoes, a solvent/solvent mixture whose function is to deliver the AI/AIs on the skin and vaporize after application. Additives are usually perfumes.

6.1.1 Computer-aided stage (S1-D)

In this stage computer-aided tools were employed in order to screen generated alternatives and propose a base case formula on which to plan experiments (stage 2) that will verify the product (stage 3).

6.1.1.1 Task S1-D1: problem definition

Sub-task S1-D1.1: performance criteria

From the knowledge base it resulted that consumers want a product which shows:

1. High effectiveness against mosquitoes, the main function of the product,
2. High compatibility with other materials (fabrics, plastics, etc),
3. Water-based, for safety and toxicology issues,
4. Good sensorial factors and cosmetic properties, that is, nice odour, appearance and good skin feeling,
5. Low-price,
6. Long durability (it should not be needed to apply the product often during exposure to mosquitoes),
7. Low toxicity,
8. High stability (no separation of phases),
9. Good user friendliness, such as a spray product,
10. Long shelf life.

Sub-task S1-D1.2: target properties

According to the knowledge base, the target properties/choices affecting the above performance criteria are:

	Performance criteria:	Target properties:
1.	Effectiveness	choice of the AI/AIs
2.	Material compatibility	choice of the solvent database (for MIXD)
3.	Water-based	inclusion of water in the solvent database (for MIXD)
4.	Cosmetic properties (odour)	choice of additives
5.	Price	cost (C)
6.	Durability	evaporation time (T_{90})
7.	Toxicity	toxicity parameter (LC_{50})
8.	Stability	Hildebrand solubility parameter (δ), Gibbs energy of mixing (ΔG^{mix}, TPD)
9.	Spray-ability	kinematic viscosity (ν), density (ρ)

The other cosmetic properties and sensorial factors (except odour), together with the shelf life, had to be validated in S2 and S3.

Sub-task S1-D1.3: constraints

Consulting the knowledge base, the constraints corresponding to the target properties defined in the previous sub-task were set:

	Performance criteria:	Target properties:	Constraints:	
1.	Effectiveness	AI/AIs	no constraints	
2.	Material compatibility	solvents	no constraints	
3.	Water-based	water	no constraints	
4.	Odour	additives	no constraints	
5.	Price	C	minimized in MIXD	
6.	Durability	T_{90}	$500 \le T_{90} \le 1500$	(6.1)
7.	Toxicity	LC_{50}	$LC_{50} \ge 0.39$	(6.2)
8.	Stability	δ	$\delta_{AI} - 3 \le \delta \le \delta_{AI} + 3$	(6.3)
			$\delta_{AI} - 3 \le \delta_{add} \le \delta_{AI} + 3$	(6.4)
		ΔG^{mix}	$\frac{\Delta G^{mix}}{RT} < 0$	(6.5)
		TPD	$TPD \ge 0$	(6.6)
9.	Spray-ability	ν	$\nu \le 75$	(6.7)
		ρ (V)	$20 \le V \le 50$	(6.8)

The units of measure and the meaning of the symbols in the above equations can be found in Table 6.1.

For additional information about the setting of the constraints, see Chapter 5 (§5.1.2.1). It has to be underlined that, instead of the density, molar volume is considered for the constraint setting ($V = Mw/\rho$). In fact, volume is additive while density is not, and additivity is important when employing linear mixing rules for the estimation of mixture properties (see sub-task S1-D3.2). The cost is minimized when calculating the solvent mixture composition (see Chapter 4, §4.1.2).

The AIs used for insect repellent are usually volatile so temperature, humidity, wind, perspiration and abrasion affect longevity; usually the high losses of repellent due to evaporation are overcome with high concentration of active ingredient, leading to high absorption on the skin, with all the safety concerns this involves (Debboun, Frances and Strickman, 2006). To increase repellent longevity without increasing the active ingredient concentration the evaporation rate of the solvent mixture has to be a reasonable compromise: if the evaporation of the solvent mixture is too fast there is the risk that part of the AI evaporates too. On the other hand if the solvent mixture

evaporates too slowly, high amount of solvents and AI can be absorbed by the skin, which should to be avoided for health reasons.

Table 6.1 lists the numerical constraint values on the target properties.

Table 6.1. Target property constraints for the insect repellent case study. UoM stands for Unit of Measure. LB and UB are the lower and the upper bound, respectively.

target property	symbol	UoM	LB	UB
evaporation time	T_{90}	s	500	1500
lethal concentration	LC_{50}	mol/l	0.39	$+\infty$
solvent mixture solubility parameter	δ	$Mpa^{½}$	$\delta_{AI} - 3$	$\delta_{AI} + 3$
additives solubility parameter	δ_{add}	$Mpa^{½}$	$\delta_{AI} - 3$	$\delta_{AI} + 3$
Δ Gibbs energy of mixing	$\Delta G^{mix}/RT$	-	$-\infty$	0
tangent plane distance	TPD	-	0	$+\infty$
kinematic viscosity	v	cS	0	75.0*
molar volume	V	l/kmol	20.0	50.0

*Bagajewicz, 2007

6.1.1.2 Task S1-D2: AI identification

Sub-task S1-D2.1: product functions

Only one AI is necessary, since the main function of the product is only one: to repel mosquitoes.

Sub-task S1-D2.2: AIs selection

Three chemicals were retrieved from the database of AIs for insect repellents:

1. DEET: it is the Active Ingredient traditionally used in insect repellents, due to its high efficiency and durability. DEET has been shown to be aggressive on clothes, plastics, glasses, and to have a high potential to irritate eyes and skin (Badolo *et al.*, 2004). It has also been blamed to be sticky, greasy and with an unpleasant odour. Finally, it causes systemic and local toxicities (Qui *et al.*, 1997);
2. Natural AIs, such as essential oils from plants (citronella, camphor, paraffin): they are safe but have limited duration (Debboun, Frances and Strickman, 2006);
3. Picaridin: it has been recently discovered. It is far superior to DEET in terms of safety, toxicology, material compatibility and cosmetic properties (Badolo *et al.*, 2004).

Picaridin (in Fig. 6.1) was chosen as the AI for this insect repellent lotion.

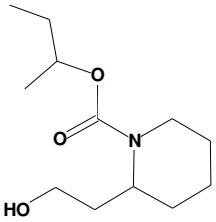


Figure 6.1. Chemical structure of picaridin, the AI chosen for the insect repellent lotion.

Sub-task S1-D2.3: AIs properties

Picaridin properties were retrieved from the database, or calculated with the property models in the model library when not available:

- Solubility information: information about picaridin solubility in water are scarce and also confused. According to Bayer (Autan®, www.autan.co.uk), picaridin has a very low solubility in water (8.6 g/l at 20° C). It has to be reminded that the objective of this case study is to design an insect repellent lotion in which water is one of the ingredients. Consequently, it has to be ensured that picaridin has a high compatibility (solubility) at least with the other solvent that will be added to water in the solvent mixture. Picaridin has a good solubility in alcohols such as ethanol and 2-propanol (www.autan.co.uk);
- Hildebrand solubility parameter (δ_{AI}): 24.1 $Mpa^{½}$, calculated with the M&G method.

Due to the low water solubility of picaridin, the constraint on the solubility parameter of the solvent mixture (Eq. 6.3) was substituted by the constraint of Eq. 6.9.

$$\delta_{AI} - 3.0 \le \delta_2 \le \delta_{AI} + 3.0 \qquad 21.1 \le \delta_2 \le 27.1 \tag{6.9}$$

δ_2 is the solubility parameter of the second solvent present in the solvent mixture (besides water). Eq 6.9 has the following meaning: only the solvents (besides water) in the solvent mixture which have a solubility parameter close to the one of picardin are able to dissolve it.
Eq. 6.4 can now be made explicit:

$$\delta_{AI} - 3.0 \le \delta_{add} \le \delta_{AI} + 3.0 \qquad 21.1 \le \delta_{add} \le 27.1 \tag{6.10}$$

6.1.1.3 Task S1-D3: solvent mixture design

Sub-task S1-D3.1: solvent database

Picaridin is very soluble in alcohols. In addition, the product under development has to contain water. Hence, the database of water soluble alcohols was retrieved from the database library and water was added, too.

Sub-task S1-D3.2: modelling choices

The mixture property models selected from the model library (see Chapter 3) for the calculation of target properties are listed in Table 6.2. The temperature considered in the design was 300 K.

Table 6.2. Models selected for the calculation of the mixture target properties for the insect repellent lotion.

target property	symbol	mixture model
evaporation time	T_{90}	Klein *et al.* (1992)
lethal concentration	LC_{50}	linear mixing rule
solubility parameters	δ, δ_{add}	linear mixing rule
Δ Gibbs energy of mixing	$\Delta G^{mix}/RT$	UNIFAC-LLE
tangent plane distance	TPD	UNIFAC-LLE
kinematic viscosity	ν	linear mixing rule
molar volume	V	linear mixing rule

Sub-task S1-D3.3: MIXD

The MIXD algorithm was applied for the property constraints of Eqs. 6.1-6.8 (Eqs. 6.3-6.4 excluded):

- Linear design level: constraints of Eqs. 6.2, 6.7 and 6.8 were applied; from a total number of 2775 binary mixtures resulting from the combination of the 75 solvents in the database (74 alcohols plus water), 2766 mixtures were rejected. Only 9 mixtures matched the constraints;
- Non-linear design level: the constraint of Eq. 6.1 was applied; no mixtures were rejected since they all matched the non-linear constraint;
- Stability test level: constraints of Eqs. 6.5-6.6 were applied; three mixtures were rejected.

Results are shown in Table 6.3. The mixtures are listed for increasing cost values. Also information about the phase stability are shown, in the last column. When a mixture shows partial miscibility (at 300 K), the compositions of the first compound in the two liquid phases are reported.
Mixtures 4, 6 and 7 are the mixtures which were rejected at the level where the stability test was performed.

Table 6.3. Mixtures matching the target properties, their property values and stability information for the insect repellent case study.

n°	mixtures	x_1	δ	v	ρ	LC_{50}	T_{90}	cost	phase
			MPa^{½}	cS	kg/l	mol/l	s	$/kg	split
1	methanol + water	0.32	0.42	0.83	0.89	0.74	818.7	0.65	stable
2	2-propanol + water	0.24	0.42	1.31	0.87	0.52	660.5	0.92	stable
3	allyl alcohol + water	0.29	0.42	1.14	0.96	0.52	598.0	1.10	stable
4	tert-butyl alcohol + water	0.24	0.42	1.49	0.94	0.45	588.2	1.22	0.02-0.44
5	ethanol + water	0.27	0.42	1.01	0.89	0.58	734.4	1.42	stable
6	2-methyl-1-propanol + water	0.23	0.42	1.66	0.88	0.42	597.0	1.72	0.02-0.46
7	2-butanol + water	0.24	0.42	1.62	0.88	0.41	519.8	1.81	0.02-0.46
8	1-propanol + water	0.25	0.42	1.28	0.88	0.47	628.2	2.07	stable

Sub-task S1-D3.4: verification

The mixture classification algorithm was applied. Mixtures are all of the type PAS/PAS, as shown in Table 6.4 (water and alcohols are polar associating fluid, PAS). Verification with rigorous models was therefore necessary.

Since viscosity is an important target property for the product being designed, viscosity was recalculated using the rigorous model of Cao *et al.* (1993). Table 6.4 compares the results of the prediction with linear and rigorous models.

A good agreement between predictions from the linear and non-linear (rigorous) models is noted and the constraint on the viscosity was not violated. R^2 is the square residual, *RD* is the percentage Relative Deviation, *SD* is the standard deviation and *AAD*(%) is the average absolute error (see Chapter 3 for their definition). In this case, the *RD* refers to the deviation of property values predicted with linear models (ζ_i^{lin}) and rigorous models (ζ_i^{rig}):

$$RD_i(\%) = \frac{\left|\zeta_i^{rig} - \zeta_i^{lin}\right|}{\zeta_i^{rig}} \cdot 100 \tag{6.11}$$

Table 6.4. Results from the verification step for the insect repellent case study. 'HB' stands for hydrogen-bonding classification. 'Lin' stands for 'linear' (mixing rule model) and 'rig' for 'rigorous' (Cao *et al.*, 1993).

n°	x_1	H-B	v-lin	v-rig	R^2	RD(%)
1	0.32	PAS/PAS	0.83	0.81	0.00	2.63
2	0.24	PAS/PAS	1.01	0.97	0.00	4.60
3	0.29	PAS/PAS	1.28	1.30	0.00	1.04
5	0.27	PAS/PAS	1.31	1.33	0.00	1.57
8	0.25	PAS/PAS	1.14	1.06	0.01	7.43
		SD/*AAD*(%)			0.04	3.46

For the five mixtures listed in Table 6.4, the constraint on the solubility parameter of the second solvent in the mixture (Eq. 6.9) was now checked. Table 6.5 lists the Hildebrand solubility parameter (δ_2) values for the alcohols of the mixtures of Table 6.4 and it can be noted that mixtures 1 and 3 had to be rejected.

Table 6.5. Solubility parameter for the alcohols involved in the mixtures reported in Table 6.3. In the last column, the matching with the constraint of Eq. 6.9 is checked.

Alcohol	mixture n°	δ_2	match Eq. 6.9?
Methanol	1	29.6	no
2-Propanol	2	23.5	yes
Allyl alcohol	3	27.5	no
Ethanol	5	26.5	yes
1-Propanol	8	24.5	yes

Sub-task S1-D3.5: optimization

The cost was the selected PI (Performance Index). Checking the cost of the remaining mixtures (2, 5 and 8) from Table 6.3, it can be noted that the mixture 2-propanol + water (mixture 2) is the cheapest therefore it was chosen as the solvent mixture for the base case product formula. Note that mixture 5 is the one used in Autan®. Autan® is composed by picaridin, a mixture ethanol + water, and a fragrance; the composition of the solvent mixture in Autan® is 28.5% (molar base) of ethanol, which is very close to the composition found in this work for the mixture ethanol + water (27% molar base). But according to the calculations performed in this work, the mixture isopropanol + water should be cheaper than ethanol + water.

Fig. 6.2 shows the reduction of candidates through the different levels of screening of sub-tasks S1-D3.3 to S1-D3.5.

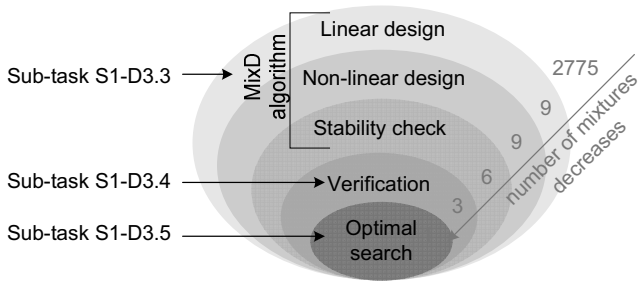


Figure 6.2. Screening of candidates in sub-tasks S1-D3.3 to S1-D3.5 for the insect repellent case study.

6.1.1.4 Task S1-D4: additives identification

Sub-task S1-D4.1: qualities to enhance

The quality to enhance was the scent.

Perfumes are added to insect repellent lotions also as fixatives for the AI, since their large branched molecules lower the vapour pressure of repellents. Tibetene and Vanillin were used as fixatives in the formulations containing DEET and they increased the longevity of 29% and 95%, respectively, when used with a ratio 1:1 with DEET (Debboun, Frances and Strickman, 2006).

Sub-task S1-D4.2: additives selection

Two aroma compounds were retrieved from the aroma database:

1. α/β-santalol;
2. Linalool.

Sub-task S1-D4.3: additives properties

The additives properties were retrieved from the aroma database, or calculated with the models in the model library when not available:

3. α/β-santalol: they are not usually found as a single aroma compound, since they are very difficult to separate. They have a sweet, woody and tenacious odour, but α-santalol shows a very rich warm-woody odour and it is the nicest between the two fragrances. They are almost colourless alcohols, slightly viscous, and with a very high boiling point (above 300 ºC). They are soluble in alcohols and oils, but insoluble in water. The solubility parameter values are: 22.34 $Mpa^{½}$ for α-santalol, 21.09 $Mpa^{½}$ for β-santalol, for an average value of 21.72 $Mpa^{½}$;
1. Linalool: it has a light and refreshing, floral woody odour with a faintly citrusy note. It is an alcohol slightly soluble in water. It has a boiling point of 198 ºC. Its solubility parameter is 21.67 $Mpa^{½}$.

Sub-task S1-D4.4: compatibility verification

Eq. 6.10 (constraint on the solubility parameter value of the additives) was here checked. Both aroma compounds match the constraint. Linalool was preferred since it is (slightly) soluble in water.

In Table 6.6 the details of the base case formula are given, along with its suggested composition, which was calculated taking as reference values in the knowledge base (Frances *et al.*, 2005). The relative amount of 2-propanol (1) and water (2) is the one listed in Table 6.3 (x_1 = 0.24).

Table 6.6. Base case formulation for the insect repellent case study. This is the formula considered in the experimental planning and verification (1st iteration).

family	chemical	Mw_i	base case formula	
			% x_i	% w_i
AI	picaridin	229.32	1.36	10.00
solvent	2-propanol	60.10	23.61	45.65
mixture	water	18.00	74.83	43.35
additives	linalool	154.24	0.21	1.00

6.1.2 Experimental planning stage (S2)

Based on the major guidelines discussed in Chapter 5, the following list of experiments was generated:

1. Measurement of the solubility limit of picaridin in water.
2. Validation of phase stability of the solvent mixture.
3. Validation of the solubility of picaridin and linalool in the solvent mixture.
4. Production of the prototype formula and validation of its phase stability.
5. ρ, v measurements for pure compounds (picaridin, isopropyl alcohol, water and linalool), solvent mixture (isopropanol + water) and formulation.
6. T_{90} measurement for the pure solvents (picaridin and linalool are high boiling and are not supposed to evaporate), solvent mixture and formulation.
7. Validation of spray-ability through a nozzle.
8. Validation of the sensorial factors and cosmetic properties: appearance (turbidity/colour), odour, stickiness, greasiness, effect on the skin, irritating power.
9. Measurement of the pH of the formula.
10. Validation of stability at different temperatures than the room temperature.
11. Validation of shelf life.

The experiments are listed according to the difficulty and time length: from the most simple and/or fast to the most difficult and/or time consuming. This is also the order which was to be employed at stage 3: experiments 1-9 were to be performed in the inner loop of stage 3 (see Fig. 5.1), while experiments 10-11 were to be performed in the outer loop (the stability test at different temperatures requires some days; the shelf life test requires at least two months of time).
Table 6.7 lists the performance criteria for the insect repellent lotion (sub-task S1-D1.1) and the corresponding experiments to be performed.

Table 6.7. Experiments employed to verify the performance criteria for the insect repellent lotion.

performance criteria	target property	considered in stage 1?	experiments planned in stage 2
effectiveness	-	yes	-[1]
material compatibility	-	yes	-
water-based	-	yes	-
good sensorial/cosmetic factors	-	only odour	Exps. 8, 9
low-priced	-	yes	-
durability	T_{90}	yes	Exp. 6
toxicity	LC_{50}	yes	-[2]
stability	δ, ΔG^{mix}, TPD	yes	Exps. 1, 2, 3, 4, 10[3]
spray-ability	v, V	yes	Exps. 5, 7
shelf life	-	no	Exp. 11

[1]no experimental facility was available
[2]usually, toxicity values are taken from MSDS in experiment-based product design
[3]experiment 10 verifies stability in a range of temperatures around 300 K, while in stage 1 only one temperature was used for the design (300 K)

The list of experiments includes also the validation of those performance criteria that were not included during the computer-aided design stage of the product. These experiments are experiments 8-11:

- Experiment 8: sensorial factors-cosmetic properties were also verified (only odour was considered during the computer-aided design stage);
- Experiment 9: the pH was measured to verify if the lotion is compatible with the skin and does not cause irritation (cosmetic products should have a pH close to that of the skin, which is 5.5);
- Experiment 10: the stability of the formula at temperatures other than 300 K (design temperature) was also tested, since transportation and storage happen at different temperatures and the product should not decompose or change appearance/odour;
- Experiment 11: the shelf life was tested, too.

The product activity could not be tested for lack of facilities. The parameter LC_{50} was not measured since in experimental works the values reported in the material safety data sheet (MSDS) for pure compounds are usually employed to calculate the mixture property value with a linear mixing rule (as done in the computer-aided design stage in this work). The solubility parameter was used in the computer-aided design stage to ensure the solubility of picaridin in the solvent mixture, but through experiments this validation was performed by observing the actual dissolution process. Therefore, solubility parameter measurements were not necessary.
The chemicals used in experiments were:

- Picaridin: from Meryer (97% w.), liquid at room temperature,
- Isopropanol: from Mallikckrodt Chemicals (minimum 99.5% w.),

- L-linalool: natural, from SAFC (minimum 80% w.),
- Water: deionised water (DI).

The experimental set-ups were:

- Solubility limit of picaridin in water: it was measured with an apparatus for liquid-liquid equilibrium. An equilibrium glass vessel (100 ml) with an external jacket was employed. Temperature was controlled through a water bath and was kept at around 293 K (20 °C) for comparison with the experimental value. A thermocouple measured the temperature of the solution picaridin-water. 100 ml of DI water was added to the vessel while picaridin was introduced in doses of approximately 0.1 g/l every day. After 3 hours mixing, the solution was left to rest for at least 18 hours, to give enough time for a possible phase separation. The experiment was concluded when phase split occurred.
- Other solubility validations: they were verified by mixing and observing if phase stability occurred. In fact, the aim of these tests was not to identify the stability limit but just to check if the designed solvent mixture was stable at the designed concentration, and if the AI and the fragrance were soluble in the mixture at the concentration of interest.
- Density measurements: a known volume of liquid was warmed up to 300 K (design temperature) with a thermal bath, and it was than weighted (density is the ratio between mass and volume).
- Viscosity measurements: a Brookfield viscosimeter (model DV-II Pro, adaptor UL/Y, spindle zero) was employed. The temperature was controlled with a thermal bath since the adaptor had a jacket.
- Evaporation time measurements: the T_{90} values (for pure compounds) employed in the computer-aided design were data (or predictions through a correlations based on such data, see Chapter 3, §3.3.1) measured with the shell thin film evaporometer according to the standard method ASTM D3539-87 (American Society for Testing and Materials, 1978). Such an apparatus was not available in the laboratories where the experiments were performed. An alternative apparatus was used for the T_{90} measurements. A qualitative filter paper (from Advantec) of 7 cm diameter was leaned on a Petri glass dish and introduced in a close precision digital scale to exclude any noises that could affect the evaporation. An amount of 0.05 ml of chemical was spread with a syringe on the filter paper, creating a circle of about 2.6 cm of diameter. The weight change was recorded along time. The percentage of weight loss was plotted against the time (seconds) and a trend was generated, from which the T_{90} could be calculated. Evaporation area is a critical parameter for evaporation phenomena, which is why it was necessary to employ a filter paper in order to achieve the same evaporation area for all the chemicals for which measurements were

performed. If the chemicals were spread directly on the Petri dish, the evaporation area would have been different due to the interaction energies with the glass. Since a different apparatus for the T_{90} measurements was employed, results could not be compared with the values predicted during the computer-aided design. But the evaporation trends gave useful information about the way the formulation evaporated and comparison between trends could be performed in order to understand how the single chemicals affected the formula evaporation.

- Spray-ability validation: this test was performed spraying the lotion though a nozzle.
- Cosmetic properties validation: greasiness, stickiness and irritating power of the formula was evaluated applying the product on the skin.
- pH measurement: pH was measured using indicator strips (Merck).
- Validation of the stability of the formula at temperatures other than 300 K: this validation was performed storing one product sample in a fridge at a temperature of 278 K (5 °C) and another sample in an oven at a temperature of 318 K (45 °C), for at least 1 week.
- Shelf life validation: a product sample was left to rest at room temperature for three months and any change in appearance, odour and consistency, as well as stability of the formula was checked.

6.1.3 Experimental validation stage (S3)

Experiments were performed in this stage to verify and amend the base case formula.

6.1.3.1 Task S3-1, iteration 1: simple/fast tests

In this task of the experimental validation stage (S3), tests 1-10 (simple tests) were performed.

Experiment 1: picaridin solubility in water

The solubility limit of picaridin in water was measured. The temperature control was oscillating between 278-281 K (20-23 °C). Phase split was noticed at 9.3 gr/l, and the value found in literature was 8.6 gr/l (@278 K). The slightly enhanced solubility can be explained with the higher temperature at which the experiment was performed.

Experiments 2-4: other solubility tests

The mixture isopropanol + water was found to be stable; picaridin, as well as linalool, could be dissolved in the solvent mixture. The product formula was found to be a single liquid phase.

Experiment 5: η and σ measurements

Tables 6.8-6.9 summarize the values used in the computer-aided design and the measured values for density, molar volume, dynamic viscosity and kinematic viscosity (the kinematic viscosity was not measured, but calculated using the experimental values of dynamic viscosity and density).

Table 6.8. Property values used in the computer-aided design (est) and values measured with experiments (exp) for the pure compounds and the solvent mixture (insect repellent case study). Units of measure: ρ [kg/m^3], V [l/kmol], η [cP], ν [cS]. 'Lin' stands for 'linear' model (linear mixing rule) and 'rig' for 'rigorous' model (viscosity calculation: Cao *et al.*, 1993).

property	picaridin		isopropanol		DI water		solvent mixture		
	exp	est	exp	est	exp	est	exp	est	
								lin	rig
ρ	1066.8	1070.0*	807.4	782.5*	965.4	1000.0*	902.7	874.8	-
V	215.0	214.3*	74.4	76.8*	18.6	18.0*	31.13	32.13	-
η	76.3	44.60	2.13	2.06*	1.02	0.89*	2.99	1.15	1.16
ν	71.5	41.68	2.64	2.63	1.06	0.89	3.31	1.31	1.33

*the value used in the calculation was experimental

Table 6.9. Experimental property values for the insect repellent case study (base case, 1st iteration).

property	1st iteration formula (base case)
ρ	952.7
V	32.62
η	3.80
ν	3.99

Properties of linalool were not measured because the amount of chemical available was not sufficient (and very expensive, like all aroma compounds).

The values used for water in the computer-aided design refer to water that was not treated with filters while the water used in the experiments was deionised water.

The estimated viscosity of picaridin is quite different from the experimental value. In fact, the M&G model employed for the estimation of the viscosity (Chapter 3; Conte *et al.*, 2008) had been shown to work very well with small molecules such as solvents but it had not been tested for multifunctional molecules with complicated structure like picaridin.

The density of the solvent mixture was predicted using a linear mixing rule (on the molar volume that is an additive property) during the computer-aided design. The linear mixing rules models are based on the assumption that there are no mixing effect when mixing water and isopropanol. Also if this assumption implies the ideality of the mixture, measurements are not far away from prediction.

The mixture viscosity was predicted both with a linear model (sub-task S1-D3.3) and the rigorous model of Cao *et al.*, 1993 (sub-task S1-D3.4). The experimental value is

reliable since it corresponds to values published by others (Pang *et al.*, 2007). The fact that the predicted value for the viscosity is far away form the experimental one is due to the difficulty in predicting the viscosity for alcohol + water mixtures, as already mentioned by Wu (1986). It has to be underlined that, however, the experimental value of the mixture viscosity still matches the viscosity constraint, and does not make the product infeasible.
Also all the other experimental values for the mixture properties match the *a priori* defined constraints. The addition of picaridin and linalool causes an increase of viscosity of less than 1 cP, since their concentration is very low compared with the concentration of water and isopropanol. This demonstrates the validity of one of the assumptions of the methodology developed in this work (see §5.1.2.2): if the AIs and additives concentrations in the formula are small, they do not significantly affect the formulation properties, which result to be very close to the solvent mixture properties.

Experiment 6: T_{90} measurements

The evaporation time measurements were performed at room temperature that oscillated between 294-295.5 K and at a humidity of 47% (approximately). Fig. 6.3 shows the trends of the percentage of weight lost during the evaporation versus the time for pure solvents, solvent mixture and overall formulation. Table 6.10 shows the T_{10}, T_{50} and T_{90} values (respectively, time at which the 10, 50 and 90% of the chemical evaporates).

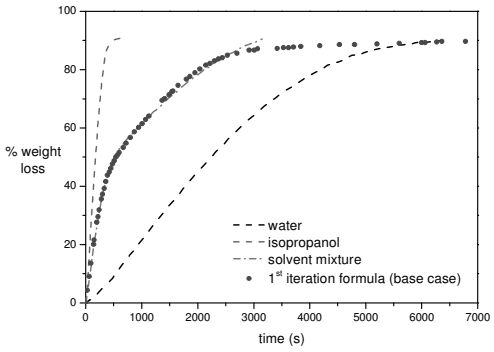


Figure 6.3. Trends of weight loss (percentage) versus time (seconds) for pure solvents, solvent mixture and 1st iteration formula (base case) for the insect repellent lotion.

Table 6.10. Evaporation times for the insect repellent case study, first experimental iteration. Evaporation times T_{10}, T_{50} and T_{90}. Unit of measure: (s).

type	T_{10}	T_{50}	T_{90}
isopropanol	32	161	505
water	384	1755	4629
solvent mixture	78	515	3082
1st iteration formula	66	540	8995

It has to be reminded that the T_{90} values can not be compared with the predicted values (see §6.1.2).
The trends of Fig. 6.3 reveal the reason why existing formulations are not based just on water or do not show high water concentration: the time for complete evaporation of water is too long. The solvent mixture trend is hidden by the formulation trend since they are almost overlapping, but while the solvent mixture reached 90% of weight loss in around 51 minutes, the overall formula reached 90% of weight loss in almost 2.5 hours.
The pure compounds show smooth trends and just before reaching the 90% of loss the trends suddenly change their slopes. Instead, the solvent mixture and the formulation trends show quite premature changes in the slopes. Both the solvent mixture and the formulation trends show a clear inflection at around 500 seconds (40-50% of weight loss): this is the point at which almost all the isopropanol had evaporated (the composition of isopropanol in the formulation is 47%). Just after almost all the isopropanol had evaporated the water started to evaporate too. At around 2700 seconds and 85-90% of weight loss there is a second inflection in the formulation trend. This is the point at which almost all the solvent mixture had evaporated. The rest of the trend reached the 90% of loss almost asymptotically, employing a very long time, since 11% (by weight) of the overall formula is constituted of picaridin and linalool, which have low vapour pressures therefore long evaporation times. After this analysis, it can be concluded that the formulation evaporated as desired: the solvent mixture evaporated at first, while the AI and the additive stayed on the desired surface for longer time, providing the desired activity, and proving a high product durability.

Experiment 7: spray-ability test

The formula could be sprayed through a commercial nozzle.

Experiments 8: validation of sensorial and cosmetic factors

The formula was completely transparent. The sensorial feeling on the skin after evaporation was a slight stickiness, caused by picaridin. No greasiness was felt, nor irritation. The scent was not pleasant since the picaridin odour was still dominant.

Experiment 9: pH measurement

The pH of the formula was measured. The pH is 8.5, too high for a personal care product (a pH between 5 and 7.5 is preferred, since the skin pH is 5.5) and it could cause irritation to some sensitive skin types.

6.1.3.2 Task S3-2, iteration 1: problems identification

The problems encountered were the unpleasant scent of the formula and the high pH value.

6.1.3.3 Task S3-3, iteration 1: amendments

The amendments suggested were:

- Increase the linalool concentration in order to improve the scent. Three prototypes were prepared, containing 2, 3 and 4% by weight of linalool, respectively. Only the 4% linalool prototype showed a satisfactory scent after the evaporation of the solvent mixture;
- Add a mild acid, such as acetic acid (glacial, 100%, from AnalaR), to correct the pH. Four prototypes were prepared with 0.05, 0.3, 0.5 and 1% by weight of acetic acid, respectively. A concentration of 0.05% rose the pH value to 5.5, which is exactly the skin pH. With an addition of 1, 0.5 or 0.3% the pH dropped down to less than 5 (too acid).

6.1.3.4 Task S3-1, iteration 2: simple/fast tests

The new product formula (2nd iteration) is given in Table 6.11.
Experiments 1 did not need to be performed again.

Table 6.11. 2nd iteration formula for the insect repellent case study. Acetic acid was added and the composition was modified from the one of the 1st iteration formula.

family	chemical	2nd iteration formula	
		$\% x_i$	$\% w_i$
AI	picaridin	1.35	9.69
solvent mixture	2-propanol	23.45	44.25
	water	74.33	42.01
additives	linalool	0.85	4.00
	acetic acid	0.03	0.05

Experiments 2-4: solubility tests

These tests were successful: after the addition of acetic acid, no phase separations were observed.

Experiments 5: η and σ measurements

The properties (density, viscosity, evaporation time) of the second iteration product formula were measured and they reported in Table 6.12. The properties did not show drastic changes since the modifications of the product formula were quite small, and they still match the *a priori* defined constraints.

Table 6.12. Property values for the 2nd iteration insect repellent lotion. Units of measure are the same as provided in Table 6.8.

property	2nd iteration formula
$\boldsymbol{\rho}$	944.7
$\boldsymbol{V}$	33.71
$\boldsymbol{\eta}$	4.27
$\boldsymbol{\nu}$	4.52

Experiment 6: T_{90} measurements

The measurement of the T_{90} was performed again on the new product formula (see in Fig. 6.4). The new formulation reached the 90% weight loss in a slightly longer time than the base case formula, due to the increased amount of linalool concentration.

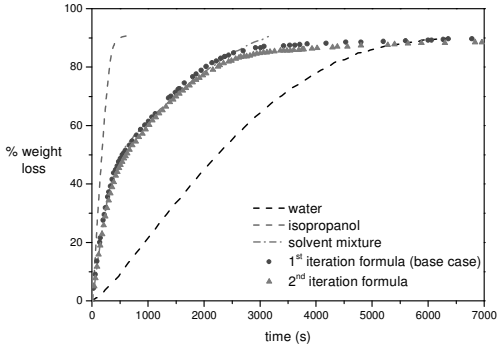


Figure 6.4. Trends of weight loss (percentage) versus time (seconds) for pure solvents, solvent mixture and 2nd iteration formula for the insect repellent lotion.

Experiment 7-9: spray-ability, sensorial-cosmetic factors and pH measurement

The product formula was still spray-able and the stickiness of the formula was reduced (this is maybe due to the lower concentration of picaridin). The pH was still 5.5 (skin pH).

6.1.3.5 Task S3-4, iteration 2: difficult tests

Since all the inner loop tests were now satisfactory, the outer loop experiments were performed.

Experiment 10: stability tests at temperatures other than 300 K

The test was satisfactory: no phase split was observed and none of the sensorial factors were affected by the temperature changes.

Experiment 11: shelf life test

Also this test was successful: after resting for two months, the product formula did not show any change.

Table 6.13 summarizes the experiments performed in each iteration, the results and amendments suggested.

Table 6.13. Experimental results for the insect repellent case study, 1st (base case) and 2nd (final) iteration.

n°	experiment	result-1st iteration	result-2nd iteration
1	solubility limit of picaridin in water	low solubility (9.3 gr/l @ 20 - 23 °C)	-
2	phase stability of the solvent mixture	successful	-
3	solubility of AI in the solvent mixture	successful	-
4	solubility of linalool in the solution picaridin + solvent mixture	successful	-
5	ρ and η of pure solvents, solvent mixture and formula	matching *a priori* defined constraints*	matching *a priori* defined constraints
6	T_{90} of pure compounds, solvent mixture and formula	satisfactory	satisfactory
7	formula spray - ability	successful	successful
8	appearance (turbidity/colour), odour, stickiness, greasiness, irritation	not satisfactory (too strong scent of picaridin), too sticky	reduced acceptable stickiness
9	pH	not satisfactory	satisfactory
10	stability at different temperatures than 300 K (278 and 318 K)	-	successful
11	shelf life	-	successful

*high deviation between predicted and experimental values for picaridin viscosity.

6.2 Case study 2: sunscreen lotion

The aim of this case study was to design a waterproof sunscreen lotion with a sun protection factor (SPF) in the range 10-15.

Sun produces a wide range of electromagnetic radiation. Ultraviolet light is responsible for sunburn and suntan, and, increases the risk of skin cancer. Ultraviolet light can be divided into: UV-A, the radiation in the 320-400 nm range; UV-B, the radiation in the 290-320 nm range; UV-C, the radiation in the 100-290 nm range. UV-C is stopped by the ozone layer in the upper atmosphere of the earth. Almost all of the UV-A and UV-B rays pass through the ozone layer and cause sunburns, skin aging and skin cancer. One of the defences of the body against UV radiation is the production of melanin, a pigment that results in darkening of the skin; but this natural defence is not enough to avoid severe damages of the skin. Sunscreens are cosmetic formulations that block UV rays.

A sunscreen lotion is constituted of AIs, a solvent/solvent mixture with the function of delivering the AIs on the skin and then evaporating, and additives which can confer better cosmetic properties (odour, skin moisture, etc).

6.2.1 Computer-aided stage

In this stage computer-aided tools were employed in order to screen numerous alternatives and propose a base case formula on which to plan experiments (stage 2) and perform experimental validation (stage 3).

6.2.1.1 Task S1-D1: problem definition

Sub-task S1-D1.1: performance criteria

From the knowledge base it resulted that consumers want a product which gives:

1. Protection from sunburns, which is one of the main functions of the product,
2. Protection from the risk of skin cancer (this requires protection against UV-A and UV-B), another main function of the product,
3. Prevention of skin aging, another main function of the product,
4. Good material compatibility,
5. Water resistance,
6. Good cosmetic properties and sensorial factors (pleasant colour and odour, pleasant skin feeling, etc),
7. Low-price,
8. Long durability, (it should not be necessary to apply it several times during the day),
9. Low toxicity,
10. Good stability,
11. User friendliness, such as a spray product,
12. Long shelf life.

Sub-task S1-D1.2: target properties

Except for the main product activities, a spray sunscreen lotion and a spray insect repellent are quite similar in terms of target properties since they both are personal care spray products:

	Performance criteria:	Target properties:
1.	Protection from sunburns	choice of the AI/AIs
2.	Protection from skin cancer	choice of the AI/AIs
3.	Prevention of skin aging	choice of the AI/AIs
4.	Material compatibility	choice of ingredients
5.	Water resistance	choice of ingredients (oil-soluble)
6.	Cosmetic properties (odour)	choice of additives
7.	Price	cost (C)
8.	Durability	evaporation time (T_{90})

9.	Toxicity	toxicity parameter (LC_{50})	
10.	Stability	Hildebrand solubility parameter (δ), Gibbs energy of mixing (ΔG^{mix}, TPD)	
11.	Spray-ability	kinematic viscosity (v), density (ρ)	

The other cosmetic properties and sensorial factors except odour, together with the shelf life, had to be validated in S2 and S3.

Sub-task S1-D1.3: constraints

Consulting the knowledge base, the constraints corresponding to the target properties defined in the previous sub-task were set:

	Performance criteria:	Target properties:	Constraints:	
1.	Protection from sunburns	AI/AIs	no constraints	
2.	Protection from skin cancer	AI/AIs	no constraints	
3.	Prevention of skin aging	AI/AIs	no constraints	
4.	Material compatibility	ingredients	no constraints	
5.	Water-resistance	ingredients	no constraints	
6.	Odour	additives	no constraints	
7.	Price	C	minimized in MIXD	
8.	Durability	T_{90}	$700 \le T_{90} \le 1300$	(6.12)
9.	Toxicity	LC_{50}	$LC_{50} \ge 3.16$	(6.13)
10.	Stability	δ	$\delta_{AI} - 3 \le \delta \le \delta_{AI} + 3$	(6.14)
			$\delta_{AI} - 3 \le \delta_{add} \le \delta_{AI} + 3$	(6.15)
		ΔG^{mix}	$\frac{\Delta G^{mix}}{RT} < 0$	(6.16)
		TPD	$TPD \ge 0$	(6.17)
11.	Spray-ability	v	$v \le 75$	(6.18)
		ρ (V)	$100 \le V \le 150$	(6.19)

Units of measure can be found in Table 6.14 (the same as in the previous case study), which also summarizes the constraint values on the target properties.

Table 6.14. Target property constraints for the sunscreen case study.

target property	symbol	UoM	LB	UB
evaporation time	T_{90}	s	700	1300
lethal concentration	LC_{50}	mol/m^3	3.16	$+\infty$
solvent mixture solubility parameter	δ	Mpa^{½}	$\delta_{AI} + 3$	$\delta_{AI} + 3$
additives solubility parameter	δ_{add}	Mpa^{½}	$\delta_{AI} + 3$	$\delta_{AI} + 3$
Δ Gibbs energy of mixing	$\Delta G^{mix}/RT$	-	$-\infty$	0
tangent plane distance	TPD	-	0	$+\infty$
kinematic viscosity	v	cS	0	75.0
molar volume	V	l/kmol	100.0	150.0

6.2.1.2 Task S1-D2: AI identification

Sub-task S1-D2.1: product functions

Three product main functions have been identified in the previous task:

- Protection from sunburns,
- Prevention of skin cancer,
- Prevention of skin aging.

Protection from sunburns and prevention of skin cancer can be achieved providing protection for both UV-A and UV-B rays. A chemical that provides screening for both types of UV radiations does not exist, therefore two different AIs have to be added. The skin aging can be prevented with antioxidants. In addition, inorganic pigments like titanium dioxide or zinc oxide are usually added to formulations, since they are opaque to light and provide a physical barrier for radiations. In conclusions, the following AIs are necessary:

- UV-A blocker,
- UV-B blocker,
- Antioxidant,
- Pigment.

Sub-task S1-D2.2: AIs selection

A list of suitable AIs has been retrieved from different databases (UV-A blockers database, UV-B blockers database, antioxidants database and pigments database) based on the following criteria:

- Oil-soluble chemicals were selected, since the product should be water resistant;
- The least toxic compounds were preferred.

The AIs chosen follow:

- UV-A blockers: 4-tert-butyl-4'-methoxydibenzoylmethane (CAS nº: 87075-14-7), well known as 'avobenzone';

- UV-B blockers: 2-ethylhexyl 2-hydroxybenzoate (CAS nº: 118-60-5), well known as 'octyl salicylate';
- Antioxidants: α-carotene, β-carotene and vitamin A;
- Inorganic pigments: titanium dioxide (TiO_2).

Fig. 6.5 shows the chemical structure of the compounds chosen as UV-filters and antioxidants for the sunscreen lotion.

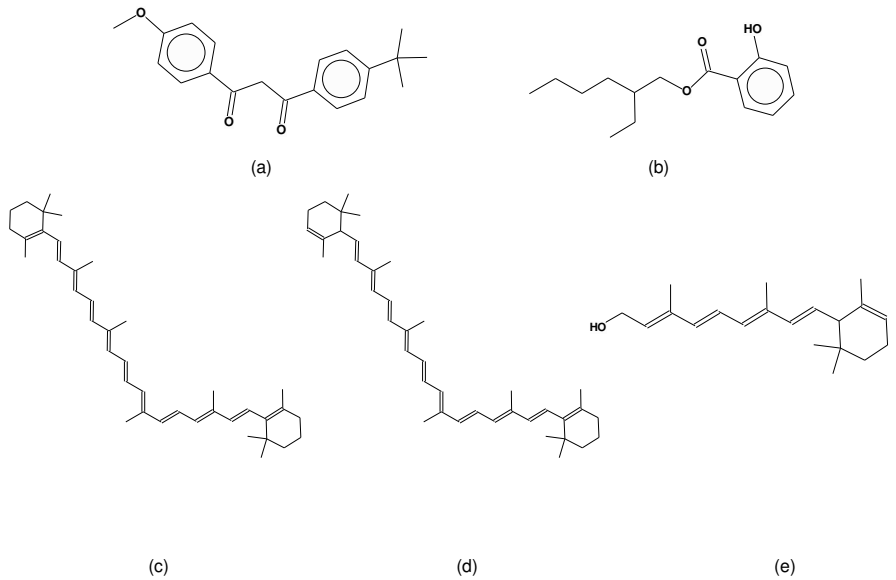


Figure 6.5. AIs considered for the sunscreen cases studies (excluded TiO_2). UV filters: (a) avobenzone and (b) octyl salicylate. Antioxidants: (c) α-carotene; (d) β-carotene and (e) vitamin A.

Sub-task S1-D2.3: AIs properties

The properties (related to solubility) of the AIs were retrieved from the databases, or calculated with property models (from the model library) when not available, and they are shown in Table 6.15. TiO_2 is not shown in Table 6.15 since it is an insoluble pigment that can be finely dispersed in the formulation, not dissolved. All the other AIs have to be dissolved in the formulation.

Table 6.15. Melting points (experimental data from databases) and solubility parameters of the AIs chosen for the sunscreen lotion (all predicted with a M&G method).

Kind of AI	AI	T_m	δ
		K	Mpa^{½}
UV filters	avobenzone	354-538	23.44
	octyl salicylate	< 298	21.50
antioxidants	α-carotene	430*	17.71
	β-carotene	451-452	17.92
	vitamin A	334-336	20.69

*predicted with a M&G method

The average value of the solubility parameters of the AIs ($\overline{\delta}_{AIs}$ = 20.3 Mpa^{½}), inorganic pigment excluded, was used to calculate the numerical values for the upper and lower bounds of Eqs. 6.14-6.15:

$$\overline{\delta}_{AIs} - 3 \le \delta \le \overline{\delta}_{AIs} + 3 \qquad 17.3 \le \delta \le 23.3 \qquad (6.20)$$

$$\overline{\delta}_{AIs} - 3 \le \delta_{add} \le \overline{\delta}_{AIs} + 3 \qquad 17.3 \le \delta_{add} \le 23.3 \qquad (6.21)$$

6.2.1.3 Task S1-D3: solvent mixture design

Sub-task S1-D3.1: solvent database

The esters database was retrieved from the database library because of the following reasons:

- The AIs selected (except TiO_2) are all oil-soluble chemicals, and esters are oils;
- The aim of the case study is to design a water resistant product;
- Esters are widely used in personal care and pharmaceuticals applications for their interesting functions (amongst which, the moisturizing effect).

Sub-task S1-D3.2: modelling choices

The models selected for the mixture target properties were the same selected for the previous case study (refer to Table 6.2). The temperature considered in the design was 300 K.

Sub-task S1-D3.3: MIXD

The MIXD algorithm was applied for all the property constraints (excluded Eq. 6.15 and therefore Eq. 6.21):

- Linear design level: constraints of Eqs. 6.13, 6.18, 6.19 and 6.20 were applied; from a total number of 4656 binary mixtures resulting from the combination of the 97 esters in the database, 4579 mixtures were rejected. Only 72 mixtures matched the linear constraints;

- Non-linear design level: the constraint of Eq. 6.12 was applied; 72 mixtures were rejected;
- Stability test level: constraints of Eqs. 6.16-6.17 were applied; no mixtures were rejected.

Results are shown in Table 6.16.
In mixtures 1, 2 and 3 the second compounds are structural isomers; the same for mixtures 4, 5 and 6. Since isomers can have property values that are very similar, the property values of the mixtures 1, 2 and 3 are close to each other, as well as the property values of the mixtures 4, 5 and 6. In addition, methoxyacetaldehyde is present in all the mixtures in high concentrations.

Table 6.16. Mixtures matching the target properties, their property values and stability information for the sunscreen case study. 'MacAl' stands for 'methoxyacetaldehyde' and 'dimethprop. 3-methbut' stands for 'dimethylpropyl 3-methylbutanoate'.

n°	mixtures	x_1	δ	v	ρ	LC_{50}	T_{90}	cost	phase split
			$Mpa^{½}$	cS	kg/l	mol/m^3	s	$/kg	
1	MacAl + 2,2-dimethylpropyl butanoate	0.89	18.95	0.53	0.83	3.63	1017.8	1.40	stable
2	MacAl + tert-butyl pentanoate	0.89	18.95	0.53	0.83	3.63	1017.8	1.40	stable
3	MacAl + isobutyl isopentanoate	0.89	18.93	0.48	0.83	3.65	878.7	1.40	stable
4	MacAl + 1,1-dimethprop. 3-methbut.	0.91	18.91	0.52	0.83	3.86	846.7	1.41	stable
5	MacAl + 2,2-dimethprop. 3-methbut.	0.91	18.92	0.53	0.83	3.80	940.2	1.41	stable
6	MacAl + isobutyl 3,3-dimethbutanoate	0.91	18.92	0.53	0.83	3.80	940.2	1.41	stable

Sub-task S1-D3.4: verification

The mixture classification algorithm was applied. Mixtures are all of the type PNA/PNA (esters are polar but non associating fluid, PNA). Verification with rigorous models was not necessary.

Sub-task S1-D3.5: optimization

The toxicity was chosen as PI. The least toxic mixture is mixture 4, methoxyacetaldehyde + 1,1-dimethylpropyl 3-methylbutanoate (the highest the value of LC_{50}, the least toxic the mixture). Fig. 6.6 shows the reduction of candidate solvent mixtures through the screening of tasks S1-D3.3 to S1-D3.5.

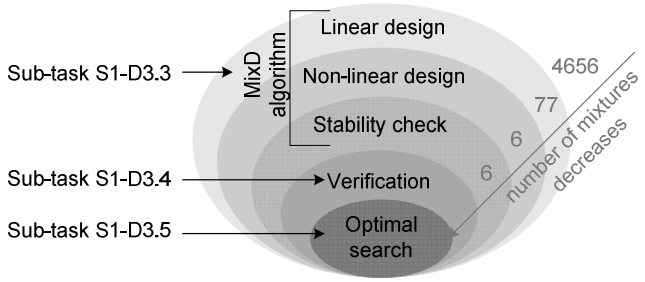


Figure 6.6. Screening of candidates in sub-tasks S1-D3.3 to S1-D3.5 for the sunscreen case study.

6.2.1.4 Task S1-D4: additives identification

Sub-task S1-D4.1: qualities to enhance

According to the performance criteria (sub-task S1-D1.1), the quality to enhance was the scent. Additional qualities to enhance/augment/enhance for a sunscreen lotion were retrieved from the knowledge base. In conclusion, the qualities to be enhanced were:

- Scent,
- UV filters protection,
- Stability,
- Microbial growth or undesirable chemical changes.

Sub-task S1-D4.2: additives selection

Chemicals that can enhance/augment/add the qualities listed in sub-task S1-D4.1 were retrieved from the additives databases:

- Enhance the scent: aroma compounds, para-menth-3-yl phenylacetate or iso-propyl salicylate;
- Augment the UV filters protection: octocrylene;
- Augment the product stability: octocrylene;
- Prevent the decomposition by microbial growth or by undesirable chemical changes: parabens.

Sub-task S1-D4.3: additives properties

The properties of the additives were retrieved from the additives databases, or calculated with the models in the model library:

1. Para-menth-3-yl phenylacetate: it is a slightly viscous ester, with a boiling point at 298 ºC and a specific gravity of 0.99; it has mild floral and extremely sweet and tenacious notes. Its solubility parameter is 17.61 $Mpa^{½}$;
2. Iso-propyl salicylate: it is an ester, with a boiling point at 240 ºC and a specific gravity of 1.08; it has a sweet ethereal herbaceous, yet quite tenacious odour with distinctly fruity character. Its solubility parameter is 24.08 $Mpa^{½}$;
3. Octocrylene: it is an ester with the ability of absorbing UV-B and short-wave UV-A rays. If used alone, it provides relatively weak screening from the sun radiations, inadequate for either UV-A or UV-B protection. This chemical is very stable and it both protects and augments the functions of other UV absorbers while improving their uniform skin coating. It solubility parameter is 18.85 $Mpa^{½}$;
4. Parabens: they are a group of esters widely used as preservatives in the cosmetic and pharmaceutical industry. Parabens are effective preservatives in many types of products. These compounds, and their salts, are used primarily for their bacteriocidal and fungicidal properties. Their efficacy as preservatives, in combination with their low cost and long history of safe use explain why parabens are so commonplace; nonyl paraben shows a solubility parameter value equal to 23.28 $Mpa^{½}$.

Sub-task S1-D4.4: compatibility verification

Para-menth-3-yl phenylacetate, iso-propyl salicylate, octocrylene and parabens, are all esters, as the AIs selected in task S1-D2 and the solvents of the mixture designed in task S1-D3.
Eq. 6.21 (constraint on the solubility parameter value of the additives) was here checked. All the additives listed in the previous sub-task match the constraint, except para-menth-3-yl phenylacetate, which is rejected. Iso-propyl salicylate is therefore selected as aroma compound for the sunscreen lotion.

In Table 6.17 the details of the base case formula are given, along with their suggested composition. The composition of the AIs in the formulation is a critical parameter since the SPF depends not only on the type of sun blocker compounds selected but also on their composition. The relation between the composition of AIs and the SPF has not been established yet. Cheng *et al.* (2009) showed that with an AIs concentration equal to 9.6% (by weight) a SPF of 6.4 is reached. The objective of this case study is to reach a SPF of 10-15, therefore a total concentration of AIs equal to 20% (by weight) is proposed.

Table 6.17. Base case formulation for the sunscreen case study. This is the formula considered in the experimental validation (1st iteration).

family	chemical	Mw_i	base case formula	
			% x_i	% w_i
AIs	Avobenzone	310.39	1.20	4.00
	Octyl salicylate	250.33	1.48	4.00
	α-Carotene	536.87	0.35	2.00
	β-Carotene	536.87	0.35	2.00
	Vitamin A	286.45	1.30	4.00
	TiO_2	79.87	4.65	4.00
solvent mixture	Methoxyacetaldehyde	74.08	83.53	66.70
	2,2-Dimethylpropyl butanoate	158.24	4.81	8.20
additives	Octorylene	361.48	0.44	1.70
	Propyl paraben	152.15	1.04	1.70
	Iso-propyl salicylate	180.20	0.88	1.70

6.2.2 Experimental planning stage (S2)

Based on the major guidelines discussed in Chapter 5, the following list of experiments was generated (it resembles the list of the previous case study):

1. Validation of the solubility of every AI in the solvent mixture.
2. Validation of solubility of every additive in the solvent mixture.
3. Production of the prototype formula and validation of its phase stability.
4. ρ, v measurement for pure solvents, solvent mixture and formulation.
5. T_{90} measurement for pure solvents, solvent mixture and formulation.
6. Validation of spray-ability through a nozzle.
7. Validation of the sensorial factors and cosmetic properties: appearance (turbidity/colour), odour, stickiness, greasiness, irritating power, soothing effect.
8. Validation of the pH of the formula.
9. Validation of stability at different temperatures than the room temperature.
10. Validation of shelf life.
11. Validation of SPF.

The experiments are listed according to the difficulty and time length: from the most simple and fast to the more difficult and time consuming. Table 6.18 lists the performance criteria for the insect repellent lotion (sub-task S1-D1.1) and the corresponding experiments to be performed.

Table 6.18. Experiments employed to verify the performance criteria for the sunscreen lotion.

performance criteria	target property	considered in stage 1?	experiments planned in stage 2
protection from sunburns	-	yes	Exp. 11
prevention of skin cancer	-	yes	Exp. 11
prevention of skin aging	-	yes	-
material compatibility	-	yes	-
water resistance	-	yes	-
good sensorial/cosmetic factors	-	only odour	Exps. 7, 8
low-priced	-	yes	-
durability	T_{90}	yes	Exp. 5
toxicity	LC_{50}	yes	-[1]
stability	δ, ΔG^{mix}, *TPD*	yes	Exps. 1, 2, 3, 9[2]
spray-ability	v, V	yes	Exps. 4, 6
shelf life	-	no	Exp. 10

[1]usually, toxicity values are taken from Material Safety Data Sheet in experiment-based product design
[2]experiments 10 verify stability in a range of temperature around 300 K, while in stage 1 only one temperature was used for the design (300 K)

The list of experiments includes also the validation of those performance criteria that were not included during the computer-aided design stage (experiments 7-10).
The chemicals used in experiments were:

- Methoxyacetaldheyde: it was not possible to obtain this chemical, therefore alternative solvent/solvent mixture needed to be considered. All the other mixtures in Table 6.16 could not be considered as alternatives, since methoxyacetaldheyde appears in all of them. The opportunity of replacing the solvent mixture was investigated going back to stage 1 (computer-aided design) following the iterative loop shown in Fig. 5.1. At task S3-D3 (mixture design task) the constraint on the toxicity parameter LC_{50} was relaxed ($LC_{50} > 0.31$ mol/m^3) and a new simulation was performed. Butyl acetate fulfilled the requirements, and it was chosen since it was available in stock.
- TiO_2: only TiO_2 nano powder was available. Due to the well known toxic (for humans) properties of small size powders, it was decided to use an available suspension of ZnO as AI (ZnO is also used as physical blockers for the UV-radiations in many sunscreen products). This suspension consists of inorganic pigment dispersed in capric/caprylic trygliceride. The solubility of the capric/caprylic trygliceride in the base case formulation of Table 6.17 (TiO_2 is not part of the formulation any more) was checked following the iterative loop of Fig. 5.1 that leads from stage 2 back to stage 1 (task S1-D2.3) of the methodology. The available property models were employed to estimate the trygliceride solubility parameter (17.72 Mpa^{½}), which was found to be very close to the solubility parameters of the other AIs (see Table 6.15).Therefore it was decided to use this ZnO dispersion for the experimental validation.

- α and β-carotene: they have very similar properties, therefore only β-carotene was purchased (from Wako).
- Iso-propyl salicylate: it was not purchased since linalool (from the previous case study) was still available. Linalool solubility parameter (21.67 Mpa½) matches the constraint on the solubility parameter (Eq. 6.21), it is therefore compatible with the other ingredients.
- Nonyl paraben: it was not found in the market. Propyl paraben (from Sigma-Aldrich) was instead purchased.
- Avobenzone: 98% w., from Meryer.
- Octyl salicylate: > 99% w. from SAFC.
- Vitamin A: natural, from H^2EI. According to the information collected about the form of the AIs (Table 6.15), Vitamin A should be solid. Instead, vitamin A from H^2EI is liquid.
- Butyl acetate: > 99.5% w., from Sigma-Aldrich.
- Octocrylene: from Meryer.

After all these considerations, the base case formula of Table 6.17 changed as shown in Table 6.19, where it is defined as '1st iteration formula'.

Table 6.19. Base case formula compared with the 1st iteration formula for the sunscreen case study.

family	chemical	composition % w_i	
		base case	1st iteration formula
AIs	avobenzone	4.0	4.0
	octyl salicylate	4.0	4.0
	α-Carotene	2.0	-
	β-Carotene	2.0	4.0
	vitamin A	4.0	4.0
	TiO_2	4.0	-
	$40\%_w$ ZnO dispersion	-	10.0 (4.0% ZnO)
solvent mixture	methoxyacetaldehyde	66.7	-
	2,2-dimethylpropyl butanoate	8.2	-
	butyl acetate	-	69.0
additives	octocrylene	1.7	1.7
	propyl paraben	1.7	1.7
	iso-propyl salicylate	1.7	-
	linalool	-	1.7

The experimental set-ups employed in this case study were the same as the insect repellent example, except for the solubility tests and the SPF test (which was not performed for case study 1):

- Solubility test: since the sunscreen lotion involves numerous AIs and additives, solubility tests were performed separately for each AI and additive in the solvent

mixture (in this case pure solvent, butyl acetate) in order to identify which chemicals cause miscibility problems. The AI (or additive) concentration in butyl acetate ($w_{AI}\big|_{AI+but.ac.}$) for the solubility tests was calculated as follows:

$$w_{AI}\big|_{AI+but.ac.} = \frac{w_{AI}\big|_{formula}}{\left(w_{AI} + w_{but.ac.}\right)\big|_{formula}} \tag{6.22}$$

w_{AI} and $w_{but.ac.}$ are the concentration of Active Ingredient and butyl acetate, respectively. The subscript 'AI+but.ac.' stands for 'solution of one AI in the solvent', while the subscript 'formula' means the 1st iteration formula of Table 6.19. When adding a solid to a liquid it is recommended to run the stirring at a higher temperature (around 40-50 °C) than the room temperature, in order to promote the dissolution process.

- The Sun Protection Factor (SPF) test followed the guidelines of FDA (Food and Drug Administration, 1978). An artificial source of light was employed: a solar simulator (Oriel #96000, 150-W) with a total power at the exit port of 8.806 mWm^{-2} (UVA: 6.906 mWm^{-2}; UVB: 1.900 mWm^{-2}) was used. The test site area was the inner part of the forearm, divided into 5 test sub site areas of 2.5 cm diameter (each). Each sub site area within a test site area was subjected for a time interval to the artificial light source for the determination of the Minimal Erythematic Dose (MED), for a series of time intervals. The rest of the skin around the sub site area was covered. At first, the MED for the unprotected skin (US) was measured in one test site area with the following time interval series: 60, 75, 94, 118, 146 seconds (following the geometric series 1.25^n). The time interval series for the protected skin (PS) test was selected in this way: the MED on unprotected skin was multiplied for the supposed Sun Protection Factor (SPF) and this time constituted the central time interval of the geometric series (if the SPF is supposed to be 4 and a subject shows an MED-US of 1.56, the time intervals to be selected for the MED-PS test are: 4, 5, 6.24, 7.84, 9.76 minutes). The SPF corresponds to:

$$\text{SPF} = \frac{\text{Exposure time interval MED}\left(\text{PS}\right)}{\text{Exposure time interval MED}\left(\text{US}\right)} \tag{6.23}$$

Uncertainty of the test is related to the interpretation of test results and depends on the individual perception of the minimal erythematic dose response (the readings can vary of ±20%, different reaction to UV light radiations in different people). Only one volunteer was used for the test; more subjects should have been used for the SPF determination, therefore the value obtained is just indicative.

6.2.3 Experimental validation stage (S3)

Experiments were performed in this stage to verify and amend the base case formula.

6.2.3.1 Task S3-1, iteration 1: simple/fast tests

In this task of the experimental validation stage (S3), tests 1-8 (simple tests) were performed.

Experiment 1-2: solubility tests on the AIs and additives

The solubility of the AIs and additives in butyl acetate was verified. The composition of the solutions produced for the solubility tests (calculated as in Eq. 6.22) and the test results are shown in Table 6.20.

Table 6.20. Concentration of the solutions AI/additive in butyl acetate for the solubility tests of the sunscreen lotion. Test results are also shown.

AI	% w_i	result
avobenzone	5.5	successful
octyl salicylate	5.5	successful
β-Carotene	5.5	failed
vitamin A	5.5	successful
ZnO dispersion	12.7	successful
octocrylene	2.4	successful
propyl paraben	2.4	successful
linalool	2.4	successful

6.2.3.2 Task S3-2, iteration 1: problems identification

β-Carotene is not soluble in butyl acetate.

6.2.3.3 Task S3-3, iteration 1: amendments

Since only one Active Ingredient out of five was found to have miscibility problems with the solvent, it was decided to substitute it with another ingredient (instead of modifying the solvent mixture, pure solvent in this case study). Therefore the external loop of Fig. 5.1 for not suitable ingredients (linking task S3-D3 with actions in stage 1) was followed. Antioxidants databases were searched and vitamin E acetate was chosen as a replacement of β-Carotene. Vitamin E acetate is a form of powdered vitamin E that is naturally converted by the body to vitamin E. It is an ester and its solubility parameter (16.91 $MPa^{½}$) is close enough to the solubility parameters of the other AIs. Vitamin E acetate was purchased from Opal.
The composition of the 2nd iteration formula is shown in Table 6.21, and compared with the base case and the 1st iteration product formulas.

Table 6.21. 2nd iteration formula compared with the base case and 1st iteration formula for the sunscreen case study.

family	chemical	composition % w_i		
		base case	1st iteration formula	2nd iteration formula
AIs	avobenzone	4.0	4.0	4.0
	octyl salicylate	4.0	4.0	4.0
	α-Carotene	2.0	-	-
	β-Carotene	2.0	4.0	-
	vitamin A	4.0	4.0	4.0
	TiO_2	4.0	-	-
	40%$_w$ ZnO dispersion	-	10.0	10.0
	vitamin E acetate			4.0
solvent mixture	methoxyacetaldehyde	66.7	-	-
	2,2-dimethylpropyl butanoate	8.2	-	-
	butyl acetate	-	69.0	69.0
additives	octocrylene	1.7	1.7	1.7
	propyl paraben	1.7	1.7	1.7
	iso-propyl salicylate	1.7	-	-
	linalool	-	1.7	1.7

6.2.3.4 Task S3-1, iteration 2: simple/fast tests

The simple and fast tests were repeated on the new formulation (2nd iteration formula), in order to verify that after the amendments suggested (§6.2.4.3), the simple and fast tests were still satisfactory.

Experiment 1: solubility tests on AIs

The solubility test on the new AI (vitamin E acetate) was performed and it was successful.

Experiment 2-3: solubility tests on additives and formulation

Experiment 2 did not need to be performed again. According to experiment 3, the prototype of the overall formulation was produced and no phase split was observed.

Experiment 4: η and σ measurements

No measurements could be performed for the solid ingredients and additives. Tables 6.22-6.23 show the calculated and experimental values for density and viscosity.
There is a big disagreement between the predicted and the measured values for the viscosity of the active ingredients, like for picaridin in the previous case study, while there is a good agreement between predictions and measurements for the solvent. This shows once again that the M&G GC$^+$ model for the prediction of the viscosity finds difficulties in predicting the properties of big and multifunctional molecules such as the AIs molecules. But these disagreements do not make the formula infeasible: molar volume and viscosity of the formulation still match the *a priori* defined targets.

Once again, the validity of one of the assumptions of the methodology developed in this work (see §5.1.2.2) is demonstrated: if the AIs and additive concentrations in the product formula are small, they do not significantly affect the formulation properties, which result very close to the solvent mixture properties.

Table 6.22. Property values used in the computer-aided design (est) and values measured with experiments (exp) for the pure compounds and the solvent mixture (sunscreen case study, 2nd iteration). Units of measure: ρ [kg/m^3], V [l/kmol], η [cP], ν [cS].

property	butyl acetate		octyl salicylate		vitamin A	
	exp	est	exp	est	exp	est
ρ	929.4	919.6	1052.1	1014.0*	978.4	934.9
V	125.0	126.3	237.9	246.9	292.8	306.4
η	0.85	0.65	9.60	92.86	91.70	12.13
ν	0.91	0.71	9.12	91.58	93.72	12.97

*the value used in the calculation is experimental

Table 6.23. Experimental property values for the sunscreen case study (2nd iteration).

property	2nd iteration formula
ρ	995.4
V	136.9
η	3.20
ν	3.21

Experiment 5: T_{90} measurements

The weight loss during evaporation (Fig. 6.7) for the solvent and the formulation was measured at the following conditions: 294-295.5 K, 47% of humidity.

The formulation trend did not reach the 90% of evaporation and a sharp change in the slope of the formulation trend could be observed between 500 and 800 seconds. After all the butyl acetate had evaporated (69% of weight loss), no evaporation was observed any more since 31% (by weight) of the formulation is composed by AIs and additives that do not evaporate in a relatively short time due to their low vapour pressure. The evaporation trend of the formulation revealed that the formula behaved as desired: the solvent evaporated almost completely after AIs and additives had been delivered on the surface (skin), providing for the desired activities.

The difference in the initial slopes of the solvent trend and the formulation trend finds justification in the fact that the evaporation rate of the solvent in the formula was slowed down by the presence of the components with low vapour pressure.

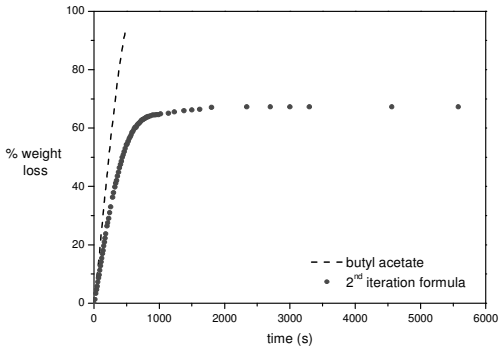


Figure 6.7. Trends of weight loss (percentage) versus time (seconds) trends for butyl acetate and 2nd iteration formula for the sunscreen case study.

Experiment 6: spray-ability test

This experiment is successful since the formula could be sprayed through a commercial nozzle.

Experiment 7: validation of sensorial and cosmetic factors

The product had the appearance of a milk since it was white and opaque due to the presence of the inorganic pigment. The scent of the overall formula was not satisfactory. When the product was sprayed on the skin, the smell of butyl acetate (sweet and fruity) was very strong but after evaporation (less than 4 minutes) no residual smell of the solvent was left on the skin, and the pleasant scent of linalool could be noticed. The sensation on the skin after spraying was good, no stickiness or greasiness was detected and the skin did not show any irritation. The soothing capacity could actually be improved.

6.2.3.5 Task S3-2, iteration 2: problems identification

The main problem was the scent of the formula, related to the solvent. But it was decided not to focus on this issue since the solvent was chosen because of its availability rather than for its characteristics. The improvement of the soothing effect was instead taken into consideration.

6.2.3.6 Task S3-3, iteration 2: amendments

In order to improve the soothing effect, an emollient needed to be added to the formulation. The outer loop of Fig. 5.1 that connects the experimental validation with the computer-aided design stage (S1) was followed, the database of moisturizing agents was searched and almond oil was found to be a feasible candidate. This oil is a blend of different fatty acids (oleic, linoleic, palmitic and stearic acid) and is well

know as emollient in personal care products. Its Hildebrand solubility parameter was calculated as an average of the single fatty acid parameters weighted on their molar fractions, and it was found to be 17.61 $Mpa^{½}$, which matches the constraint of Eq. 6.21. Usually, for an emollient, a concentration between 2-5% w. is employed. Two prototypes were produced, one with a concentration of 2% and another with a concentration of 4%. It was found that a concentration of 2% was sufficient to give the desired effect on the skin. The composition of the new formula (3rd iteration formula) is shown in Table 6.24 and compared with the previous formulations.

Table 6.24. 3rd iteration formula compared with the 1st and 2nd iteration formulas for the sunscreen case study.

family	chemical	composition % w_i		
		1st iteration formula	**2nd iteration formula**	**3rd iteration formula**
AIs	avobenzone	4.0	4.0	4.0
	octyl salicylate	4.0	4.0	4.0
	α-Carotene	-	-	-
	β-Carotene	4.0	-	-
	vitamin A	4.0	4.0	4.0
	TiO_2	-	-	-
	$40\%_w$ ZnO dispersion	10.0	10.0	10.0
	vitamin E acetate		4.0	4.0
solvent mixture	methoxyacetaldehyde	-	-	-
	2,2-dimethylpropyl butanoate	-	-	-
	butyl acetate	69.0	69.0	67.0
additives	octocrylene	1.7	1.7	1.7
	propyl paraben	1.7	1.7	1.7
	iso-propyl salicylate	-	-	-
	linalool	1.7	1.7	1.7
	almond oil	-	-	2.0

6.2.3.7 Task S3-1, iteration 3: simple/fast tests

Experiment 1 did not need to be performed again.

Experiment 2-3: solubility tests on additives and formulation

It was verified that almond oil is soluble in butyl acetate. In addition, the 3rd iteration formula prototype was produced and almond oil was found to be compatible with the other ingredients (no phase split was observed).

Experiment 4: η and σ measurements

Since almond oil looked very viscous, its viscosity was measured in order to understand the impact of an addition of 2% w. to the formulation. The dynamic viscosity resulted to be 61.70 cP (ν = 65.9 cS). The properties of the formula at the 3rd iteration loop were measured too. Table 6.25 shows the experimental values.

Table 6.25. Experimental property values for the sunscreen case study (3rd iteration). The units of measure are the same as provided in Table 6.22.

property	3rd iteration formula
ρ	1038.8
V	133.0
η	3.20
ν	3.08

Even if almond oil is highly viscous, a concentration of 2% by weight in the overall formula does not affect the formula viscosity. The formula property values still match the constraints.

Experiment 5: T_{90} measurements

The evaporation time of the 3rd iteration formula was measured (Fig. 6.8).

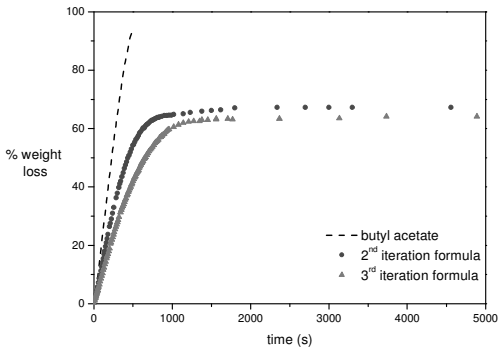


Figure 6.8. Trends of weight loss (percentage) versus time (seconds) for butyl acetate, 2nd and 3rd iteration formulas for the sunscreen case study.

The evaporation trend is slightly different from the previous formula trend, due to the addition of an extra compound with low vapour pressure. As a consequence, the evaporation was further slowed down at the beginning and the asymptotic value of weight loss reached was even lower than the one reached by the 2nd iteration formula (the total amount of AIs and additives in the formula increased from 31% to 33%).

Experiment 6-8: spray-ability, sensorial-cosmetic factors, pH measurement

The product is spray-able. The formula appeared white and opaque and with the same strong scent as before. The product was not sticky or greasy and it did not cause irritation. The pH of the formula was still 5.5 (skin pH).

6.2.3.8 Task S3-4, iteration 3: difficult tests

The experimental iterative loop for the simple and fast tests was successfully completed (except for the scent of the formula). Now the difficult and time consuming experiments could be performed.

Experiment 9: stability tests at temperatures other than 300 K

At first, the stability test at different temperatures was performed. Deposition of the zinc oxide was observed at 45 °C, along with a mild change in the colour of the formula, which appeared slightly yellow after the test. It is important that all the inorganic pigment is dispersed in the lotion to ensure the specified sun protection factor. The inorganic pigment could be dispersed with a brief shake of the sample.
No changes were observed for the sample of the test performed at 5 °C.

Experiment 10: shelf life test

The shelf life test revealed a partial deposition of ZnO. Also in this case, the inorganic pigment could be dispersed with a brief shake of the sample. The shelf life could be considered satisfactory under the condition the product is shaken before use (and this condition has to be specified upfront, for instance, on the product container).

Experiment 11: SPF validation

The SPF test was now performed. The MED-US (unprotected skin) for the volunteer was recorded at 118 seconds. The sunscreen lotion was estimated to have a SPF around 8, to be precautionary. Therefore erythema on the volunteer should have be recorded at 8·118 seconds, which corresponds to 944 seconds (15 minutes and 44 seconds). Therefore the time sequence chosen for the test was: 531, 708, 944, 1180, 1475 seconds. The MED-PS (protected skin) was detected at 708 seconds, leading to a SPF value of 6, which does not match the target (10-15).

6.2.3.9 Task S3-3, iteration 3: problems identification

The test at 45 ºC was not completely satisfactory since the product changes colour. The SPF did not match the targets. Amendments were necessary, but further experimental work was not performed and it is not discussed here.

Table 6.26 summarizes the experimental iterations, the tests performed and the results obtained.

Table 6.26. Experimental results for the sunscreen case study, 1st, 2nd and 3rd iteration.

n°	experiment	result-1st iteration	result-2nd iteration	result-3rd iteration
1	AIs solubility in the solvent mixture	one AI does not dissolve	successful	-
2	additives solubility in the solvent mixture	-	successful	successful (performed only for almond oil)
3	phase stability of the overall formula	-	successful	successful
4	ρ and η of pure compounds, solvent mixture and formula	-	matching targets*	matching targets
5	T_{90} of pure solvents, solvent mixture and formula	-	satisfactory	satisfactory
6	formula spray – ability	-	successful	satisfactory
7	appearance (turbidity/colour), odour, stickiness, greasiness, irritation, soothing effect	-	odour not satisfactory**, soothing effect could be improved	odour not satisfactory**, soothing effect improved
8	pH	-	satisfactory	satisfactory
9	stability at different temperatures than 300 K (278 and 318 K)	-	-	not satisfactory
10	shelf life	-	-	satisfactory with condition***
11	SPF	-	-	not satisfactory

*high deviations of measured viscosity values from predicted values for some AIs viscosity (octyl salicylate and vitamin A)
**the odour is not satisfactory but this factor is taken into consideration for improvements because of chemicals availability issues
***the condition is that the product has to be shaken before use

6.3 Case study 3: paint formulation

The aim of this case study was to design a white waterproof (for exterior) paint formulation, to employ for the finishing of surfaces. For this product, the stages of experimental planning (stage 2) and validation (stage 3) were not performed and therefore, results only from the computer aided design stage are reported below.
A paint formulation is constituted of a pigment conferring the particular colour, a binder with the function of binding the insoluble pigment particles and providing the surface coating, a solvent or a mixture of solvents whose function is to deliver the paint on a surface (when it is applied) and then vaporizes. Additives may be added to give the paint formulation a particular appearance or to enhance its spread-ability on surface during application and so on (Van der Walle *et al.*, 1999).

6.3.1 Computer-aided stage (S1-D)

In this stage computer-aided methods and tools were employed to generate and screen alternatives and propose a base case formula.

6.3.1.1 Task S1-D1: problem definition

Sub-task S1-D1.1: performance criteria

From the knowledge base it resulted that consumers want a product which shows:

1. Desired surface colour and coating (main function of the product); in this case study the desired colour is white,
2. Water-proofness,
3. Low-price,
4. Short drying time,
5. Low toxicity,
6. Good spread-ability on surfaces,
7. Good stability (no separation of phases which requires mixing of the product before application to make it homogeneous).

Sub-task S1-D1.2: target properties

According to the knowledge base, the target properties/choices affecting the above performance criteria are:

	Performance criteria:	Target properties:
1.	Colour/coating	choice of AI/AIs
2.	Water-proofness	choice of all ingredients (oil-soluble)
3.	Price	cost (C)
4.	Drying time	evaporation time T_{90}
5.	Toxicity	toxicity parameter (LC_{50})
6.	Spread-ability	dynamic viscosity (η), surface tension (σ), density (ρ)
7.	Stability	Hildebrand solubility parameter (δ), Gibbs energy of mixing (ΔG^{mix}, TPD)

Sub-task S1-D1.3: constraints

Consulting the knowledge base, the constraints corresponding to the target properties defined in the previous sub-task were set:

	Performance criteria:	Target properties:	Constraints:	
1.	Colour/coating	AI/AIs	no constraints	
2.	Water-proofness	ingredients	no constraints	
3.	Price	C	minimized in MIXD	
4.	Drying time	T_{90}	$255 \le T_{90} \le 450$	(6.24)
5.	Toxicity	LC_{50}	$LC_{50} \ge 0.40$	(6.25)
6.	Spray-ability	η	$0.6 \le \eta \le 0.9$	(6.26)
		σ	$26.5 \le \sigma \le 29.5$	(6.27)

		$\rho\ (V)$	$100 \le V \le 130$	(6.28)
7.	Stability	δ	$\delta_{AI} - 3 \le \delta \le \delta_{AI} + 3$	(6.29)
			$\delta_{AI} - 3 \le \delta_{add} \le \delta_{AI} + 3$	(6.30)
		ΔG^{mix}	$\frac{\Delta G^{mix}}{RT} < 0$	(6.31)
		TPD	$TPD \ge 0$	(6.32)

Regarding the drying time, it is important that the paint formulation does not dry too slowly. For safety and environmental reasons also a lower bound for the drying time is desirable.
Table 6.27 lists the target properties and the relative constraint values (upper and lower bounds) for the paint formulation. Units of measure are also reported.

Table 6.27. Target property constraints for the paint case study.

target property	symbol	UoM	LB	UB
evaporation time	T_{90}	s	255	450
lethal concentration	LC_{50}	mol/m^3	0.40	$+\infty$
dynamic viscosity	η	cP	0.6	0.9
surface tension	σ	mN/m	26.5	29.5
molar volume	V	l/kmol	100.0	130.0
solvent mixture solubility parameter	δ	Mpa^{½}	$\delta_{AI} - 3$	$\delta_{AI} + 3$
additives solubility parameter	δ_{add}	Mpa^{½}	$\delta_{AI} - 3$	$\delta_{AI} + 3$
Δ gibbs energy of mixing	$\Delta G^{mix}/RT$	-	$-\infty$	0
tangent plane distance	TPD	-	0	$+\infty$

6.3.1.2 Task S1-D2: AI identification

Sub-task S1-D2.1: product functions

The function of the paint formulation is to provide the desired colour (white) and a protective coating (water resistant) on the surface.

Sub-task S1-D2.2: AIs selection

From the knowledge base, it was found that pigments are used for coloring paint, ink, plastic, fabric, cosmetics, food and other materials. Pigment particles are insoluble, so a binder is needed to bind together and to provide the protective coating on the surface. Polymers are common binders for pigments.
From the pigments and polymers databases, the following chemicals were retrieved:

- Pigment database: titanium dioxide (TiO_2). It is a common white pigment, which is insoluble and precipitates;
- Polymers database: polyesters such as poly(3-hydroxylalkanoates) (PHAs). They constitute a family of biodegradable polymers frequently used in paints,

which are also able to bind the TiO_2 particles. In this case study the simplest polymer from the family of PHAs is considered (Fig. 6.9).

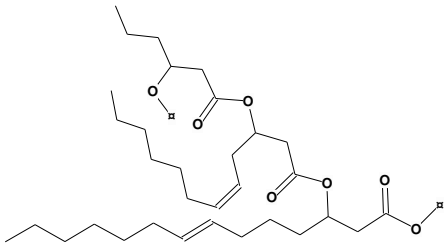


Figure 6.9. The simplest polymer from the family of the poly(3-hydroxylalkanoates), selected as one of the AIs for the paint formulation designed in this case study.

Sub-task S1-D2.3: AIs properties

The AIs properties are retrieved from the database or calculated with the models in the model library when not available:

- Solubility information: TiO_2 is insoluble in any solvents. PHAs are polyesters, therefore they are water insoluble and they can be employed for the paint formulation designed in this case study, which has to be a waterproof product;
- Hildebrand solubility parameter: the needed solubility parameter is the one of the polymer, which binds the pigment and has to be dissolved in the solvent/solvent mixture. The polymer solubility parameter is 19.92 $Mpa^{½}$.

The constraints on the mixture solubility parameter (Eq. 6.29) and the additives solubility parameters (Eq. 6.30) therefore become:

$$\delta_{Pol} - 3.0 \le \delta \le \delta_{Pol} + 3.0 \qquad 17.0 \le \delta \le 23.0 \tag{6.33}$$

$$\delta_{Pol} - 3.0 \le \delta_{add} \le \delta_{Pol} + 3.0 \qquad 17.0 \le \delta_{add} \le 23.0 \tag{6.34}$$

δ_{pol} is the solubility parameter of the polymer, which is the AI that has to be dissolved by the solvent/solvent mixture.

6.3.1.3 Task S1-D3: solvent mixture design

Sub-task S1-D3.1: solvent database

The paint formulation under development has to be water insoluble. Hence, the database of water insoluble solvents usually employed for paint formulations is retrieved from the database library.

Sub-task S1-D3.2: modelling choices

The mixture property models selected from the model library (see Chapter 3) for the calculation of target properties are shown in Table 6.28. The temperature considered is 300 K.

Table 6.28. Models selected for the calculation of the mixture target properties, paint formulation case study.

target property	symbol	mixture model
evaporation time	T_{90}	Klein *et al.* (1992)
lethal concentration	LC_{50}	linear mixing rule
dynamic viscosity	η	linear mixing rule
surface tension	σ	linear mixing rule
molar volume	V	linear mixing rule
solubility parameter	δ, δ_{add}	linear mixing rule
Δ Gibbs energy of mixing	$\Delta G^{mix}/RT$	UNIFAC-LLE
tangent plane distance	TPD	UNIFAC-LLE

Sub-task S1-D3.3: MIXD

The MIXD algorithm was applied for all the property constraints (excluded Eq. 6.30 and therefore Eq. 6.34):

- Linear design level: constraints of Eqs. 6.25-6.28 and 6.33 were applied; from a total number of 465 binary mixtures resulting from the combination of the 31 solvents in the database, 406 mixtures were rejected. Only 59 mixtures matched the linear constraints;
- Non-linear design level: the constraint of Eq. 6.24 was applied; 48 mixtures were rejected;
- Stability test level: constraints of Eqs. 6.31-6.32 were applied; no mixtures were rejected.

Results are shown in Table 6.29, where the solvent mixtures are listed in terms of increasing cost. Note that mixture 10 is the only mixture showing miscibility problems, but at the designed molar fraction of 0.90 the mixture it is a single liquid phase. Therefore this mixture was not rejected.

Table 6.29. Mixtures matching the target properties, their property values and stability information for the paint case study. 'DEGEE' stands for 'diethylene glycol ethyl ether'.

n°	mixtures	x_1	δ	η	σ	ρ	LC_{50}	T_{90}	cost	phase
			Mpa^{½}	cP	mN/m	kg/l	mol/l	s	$/kg	split
1	DEGEE + toluene	0.05	18.35	0.71	28.63	0.87	0.40	256.1	1.35	stable
2	toluene + cyclohexanone	0.95	18.24	0.63	28.80	0.86	0.40	256.4	1.37	stable
3	toluene + butyrolactone	0.96	18.49	0.60	28.65	0.86	0.43	255.6	1.60	stable
4	toluene + ethylbenzene	0.56	18.03	0.60	28.41	0.86	0.40	345.8	2.49	stable
5	ethylbenzene + heptane	0.62	17.67	0.67	26.93	0.87	0.40	448.0	3.51	stable
6	ethylbenzene + ethyl acetate	0.87	17.46	0.60	26.94	0.84	0.52	437.4	3.67	stable
7	ethylbenzene + butyl acetate	0.87	17.42	0.60	26.50	0.84	0.46	436.1	3.70	stable
8	ethylbenzene + hexane	0.77	17.92	0.60	27.18	0.87	0.69	403.8	3.72	stable
9	ethylbenzene + butanone	0.80	18.09	0.60	27.56	0.85	1.02	427.2	3.72	stable
10	ethylbenzene + dichloromethane	0.90	18.11	0.60	28.19	0.89	0.62	449.7	3.85	0.0-0.71
11	ethylbenzene + isopropylacetone	0.72	17.54	0.62	26.50	0.84	0.84	418.2	3.91	stable

Sub-task S1-D3.4: verification

The mixture classification algorithm was applied. Mixtures 4, 5 and 8 are mixtures of two normal fluids (hydrocarbons, cyclohydrocarbons and derivates). These mixtures do not need verification with rigorous models, and they were not therefore considered in this sub-task.

For all the other mixtures, the HB classification is reported in Table 6.30. Viscosity and surface tension are critical parameters for a product that has to be spread on a surface, therefore verification was performed using the model of Cao *et al.* (1993) for viscosity and the model of Suarez *et al.* (1989) for surface tension. Table 6.30 lists the results of the calculations.

Table 6.30. Results from the verification step for the paint case study. 'HB' stands for hydrogen-bonding classification. 'Lin' and 'rig' stand for 'linear' (mixing rule model) and 'rigorous' (η: Cao *et al.*, 1993; v: Suarez *et al.*, 1989). Highlighted in a different colour, the values which do not match the constraints of Eqs. 6.26-6.27.

n°	x_1	HB	η-prediction			σ-prediction		
			linear	Cao	RD(%)	linear	Suarez	RD(%)
1	0.05	PAS-NF	0.71	0.62	14.02	28.63	28.61	0.05
2	0.95	NF-PNA	0.63	0.60	5.66	28.80	28.70	0.35
3	0.96	NF-PNA	0.60	0.58	3.71	28.65	28.60	0.20
6	0.87	NF-PNA	0.60	0.59	2.10	27.00	25.48	5.98
7	0.87	NF-PNA	0.61	0.61	1.34	27.00	24.94	8.28
9	0.80	NF-PAS	0.60	0.60	0.10	27.56	27.53	0.12
10	0.90	NF-PNA	0.60	0.60	0.25	28.19	28.22	0.12
11	0.72	NF-PNA	0.63	0.63	0.76	27.00	26.58	1.59
	SD/AAD(%)				3.49			2.09

Mixtures 3 and 6 did not match the target on viscosity ($0.6 < \mu < 0.9$), while mixtures 6 and 7 did not match the target on surface tension ($26.5 < \sigma < 29.5$), therefore mixtures 3, 6 and 7 were rejected and not considered in the subsequent sub-tasks. The mixtures considered in the next sub-task were: mixture 1, 2, 4, 5, 8, 9, 10 and 11.

Sub-task S1-D3.5: optimization

Cost was selected as the only performance index PI. To determine the optimal mixture requires only a check of Table 6.29, which lists the mixture in terms of increasing cost: mixture 1 is the cheapest between mixtures 1, 2, 4, 5, 8, 9, 10 and 11. The mixture diethylene glycol ethyl ether + toluene is selected as the solvent mixture for the paint formulation.

Fig. 6.10 highlights the reduction of the number of feasible solvent mixtures through sub-tasks S1-D3.3 to S1-D3.5.

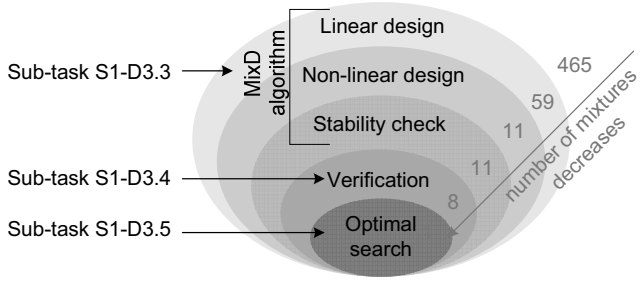


Figure 6.10. Screening of candidates in sub-tasks S1-D3.3 to S1-D3.5 for the paint case study.

6.3.1.4 Task S1-D4: additives identification

Sub-task S1-D4.1: qualities to enhance

From the knowledge base it was found that the spread-ability on surfaces is a very important quality for paint formulations. Hence, it was decided to look for additives that could enhance this property.

Sub-task S1-D4.2: additives selection

Wetting agents lower the surface tension of the blend they are added to, and they are used to enhance the spread-ability of paint formulations. Two additives were retrieved from the wetting agents database:

1. Sodium dioctyl sulfosuccinate (Aerosol OT);
2. Acetylenic surfactant (Surfynol^{TM} 104).

They are both powerful wetting agents because the presence of the polar group (in between the surfactant chains) provides sufficient spacing for the solvent to reach the surface of the pigment particle.

Sub-task S1-D4.3: additives properties

The following information were retrieved from the wetting agents database or calculated through the models contained in the model library:

1. Sodium dioctyl sulfosuccinate (Aerosol OT): it is water insoluble and has a solubility parameter of 22.95 $Mpa^{½}$. It is found that sodium dioctyl sulfosuccinate also promotes the solubilisation and the dispersion of the pigment solid particles;
2. Acetylenic surfactant (Surfynol^{TM} 104): it is biodegradable and has a solubility parameter of 23.95 $Mpa^{½}$.

Sub-task S1-D4.4: compatibility verification

The constraint of Eq. 6.34 was here applied to verify the compatibility of the selected additive with the designed solvent mixture and the selected AIs. Sodium dioctyl sulfosuccinate solubility parameter matches Eq. 6.34, while the acetylenic surfactant solubility parameter does not match Eq. 6.34. Hence, sodium dioctyl sulfosuccinate was selected a wetting agent for the paint formulation designed in this case study.

In Table 6.31 the base case formulation for a white and water insoluble paint for finishing of exteriors is shown. The suggested composition was calculated taking values from literature as reference (Tarng *et al.*, 2010).

Table 6.31. Base case formulation for the paint case study.

family	compound	Mw_i	Base case formula	
			$\% x_i$	$\% w_i$
AIs	TiO_2	79.87	6.04	3.40
	PHAs	12.48E+4	0.04	34.0
solvent mixture	DEGEE	134.17	2.96	2.80
	toluene	92.14	90.62	58.80
additives	sodium dioctyl sulfosuccinate	422.57	0.34	1.00

6.4 Case study 4: virtual PPD-lab application

This case study highlights the application of the new work-flow for formulation design and verification implemented in the virtual PPD-lab. The product being considered here is an alternative formulation for an insect repellent lotion, based just on alcohols (no water). For this example, stage 2 and stage 3 of the methodology are not highlighted since the objective is to illustrate the use of the virtual PPD-lab.
The product details are the same as given in §6.1 for the water-based insect repellent lotion, except for the fact that water does not have to appear as one of the ingredients in the final formulation since this product is alcohol-based.
Changes in the problem description were necessary:

- The constraint on the solubility parameter was defined directly on the solvent mixture, in the MIXD algorithm, and not on the alcohol present in the mixture. This constraints was therefore applied during the design of the solvent mixture (sub-task S1-D3.3) and not during the problem verification (sub-task S1-D3.4), as it was done in the previous insect repellent case study;

- Both the databases of water soluble and water insoluble alcohols were retrieved for this case study. Water was not included.

6.4.1 Computer-aided design stage (S1-D)

The virtual PPD-lab (see Appendix F) was employed in this stage (computer-aided stage).

6.4.1.1 Task S1-D1: problem definition

The performance criteria and user needs that were defined in §6.1.1.1, are also valid for the alcohol-based insect repellent lotion. Therefore sub-tasks S1-D1.1 and S1-D1.2 do not need to be described again. Table 6.32 summarizes the constraints (sub-task S1-D3.3) for this design problem.

Table 6.32. Target property constraints for the alcohol-based insect repellent lotion.

target property	symbol	UoM	LB	UB
evaporation time	T_{90}	s	500	1500
lethal concentration	LC_{50}	mol/l	0.05	$+\infty$
solvent mixture solubility parameter	δ	$Mpa^{½}$	$\delta_{AI} - 3$	$\delta_{AI} + 3$
additives solubility parameter	δ_{add}	$Mpa^{½}$	$\delta_{AI} - 3$	$\delta_{AI} + 3$
Δ Gibbs energy of mixing	$\Delta G^{mix}/RT$	-	$-\infty$	0
tangent plane distance	TPD	-	0	$+\infty$
kinematic viscosity	v	cS	0	75.0
molar volume	V	l/kmol	75.0	100.0

Solution through the virtual PPD-lab

At first, the 'Problem definition' task in the virtual PPD-lab (which corresponds to task S1-D1 in the methodology shown in Fig 5.1) was accessed, and the 'insect repellent' option was chosen from the optional products in the list (knowledge base), as shown in Fig. 6.11. In the next dialogue box (Fig. 6.12) a list of performance criteria (knowledge base) is proposed by the virtual PPD-lab, and the necessary performance criteria for an insect repellent were selected, as in sub-task S1-D1.1. The translation from user needs to target properties (knowledge base) was performed in the dialogue box of Fig. 6.13, as in sub-task S1-D1.2. This translation is a default option (knowledge base). If necessary, the user can also make his/her own selection for the target properties (see Appendix F for more information).

Constraints needed to be set, as in sub-task S1-D1.3. This was performed in the next dialogue box (Fig. 6.14). In the virtual PPD-lab the logarithm of the toxicity parameter LC_{50} is employed. Therefore the constraint on LC_{50} (given in Table 6.32) becomes:

$$0.0 < -\log_{10}\left(LC_{50}\right) < 1.3 \tag{6.35}$$

The constraint on the solvent mixture solubility parameter (upper and lower bounds) could not be set yet. In fact, this constraint requires to establish the solubility parameter of the AI first.

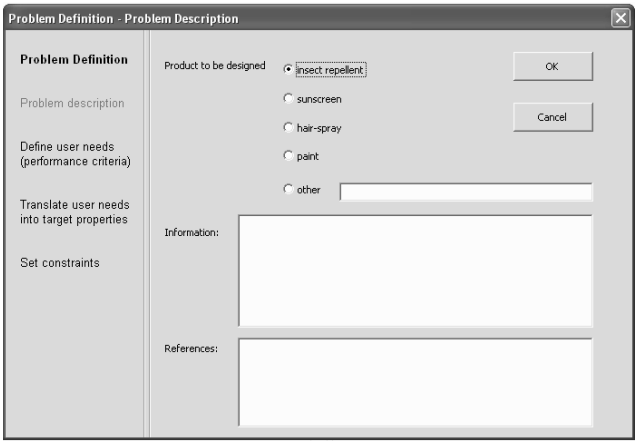


Figure 6.11. The first dialogue box of the problem definition task, where the user can select a product in the list or add a new product. The insect repellent lotion is selected.

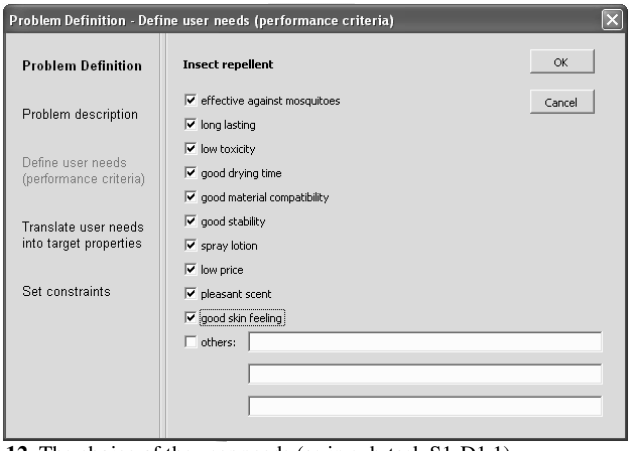


Figure 6.12. The choice of the user needs (as in sub-task S1-D1.1).

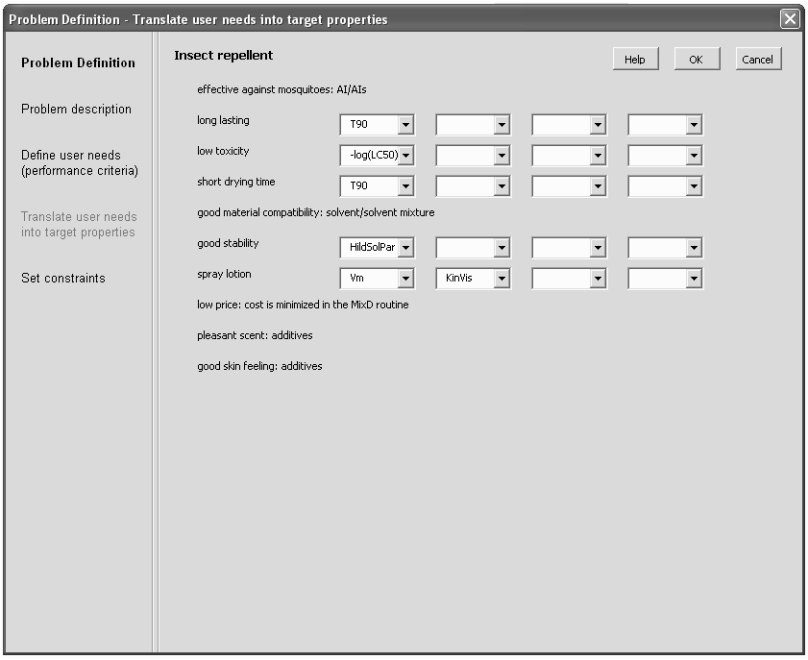


Figure 6.13. The default translation of the performance criteria into user needs (as in sub-task S1-D1.2). The meaning of the symbols for the target properties is explained in the 'Help' on the top right of the dialogue box. 'HildSolPar' stands for 'Hildebrand solubility parameter', 'Vm' for 'molar volume', 'KinVisc' for 'kinematic viscosity'.

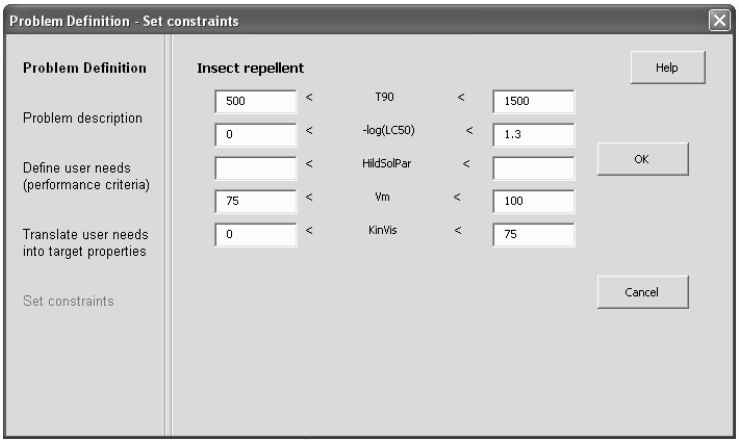


Figure 6.14. Setting of the constraint on the target properties (as in sub-task S1-D1.3).

6.4.1.2 Task S1-D2: AI/AIs identification

The function of the product is to repel mosquitoes (sub-task S1-D2.1) and the AI selected is the picaridin (sub-tasks S1-D2.2 and S1-D2.3), as in §6.1.1.2.

Solution through the virtual PPD-lab

First, the 'AI/AIs selection' task (which corresponds to task S1-D2 in the methodology shown in Fig 5.1) was accessed. The product activity was selected, as in sub-task S1-D2.1. This is shown in the dialogue box of Fig. 6.15. Then, a chemical that can provide for the selected activity was identified, as in sub-task S1-D2.2. Picaridin was chosen from the AIs database in the dialogue box of Fig. 6.16. Finally, the AI properties were collected, as in sub-task S1-D2.3. This is shown in the dialogue box of Fig. 6.17.

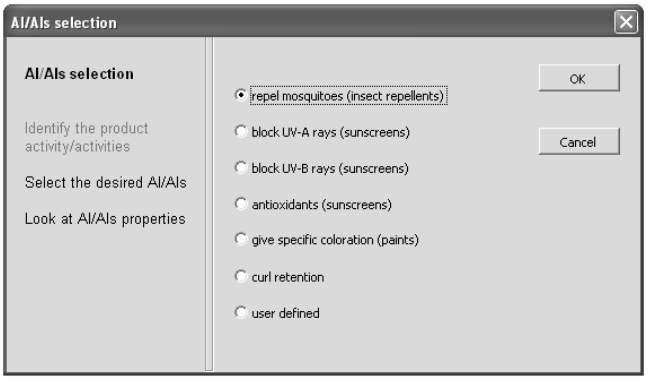


Figure 6.15. Selection of the product activity (as in sub-task S1-D2.1).

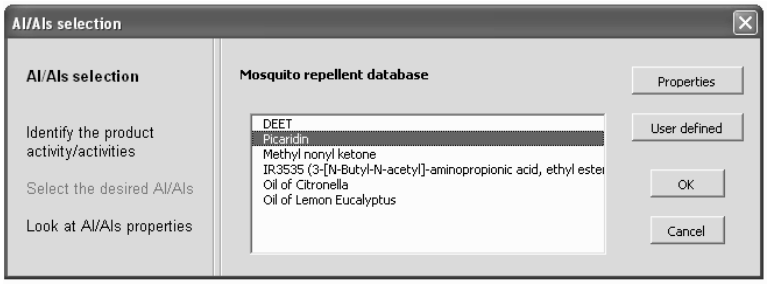


Figure 6.16. The choice of the AI (as in sub-task S1-D2.2): picaridin.

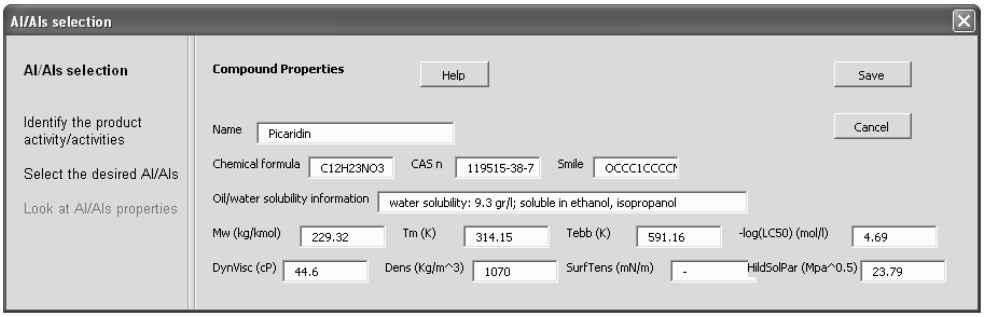


Figure 6.17. The collection of data/information for the selected AI (as in sub-task S1-D2.2).

The AI Hildebrand solubility parameter was now known (23.79 $Mpa^{1/2}$), therefore the dialogue box of Fig. 6.14 was accessed again to set the upper bound and lower bound values for the constraint on the solubility parameter.

6.4.1.3 Task S1-D3: solvent mixture design

The suitable solvents for this case study are alcohols (sub-task S1-D3.1). The modelling choices performed were the same discussed in §6.1.1.3 (sub-task S1-D3.2). The results from the MIXD program (sub-task S1-D3.3) are displayed in Tables 6.33-6.34. All the listed mixtures show complete miscibility.

The mixture classification algorithm was applied and it resulted that all the mixtures are of the type PAS/PAS. Hence, it was necessary to verify all the mixtures of Table 6.33 (sub-task S1-D3.4). The results of the verification of the viscosity with the rigorous model of Cao *et al.* (1993) are given in Table 6.35. All the mixtures still match the constraint on the kinematic viscosity.

Table 6.33. Solvent mixtures matching the a priori defined targets for the alcohol-based insect repellent lotion.

n° component names	x_1	v	V	δ	LC_{50}	T_{90}	Cost
		(cS)	(l/kmol)	($Mpa^{1/2}$)	(mol/l)	(s)	($/kg)
1 methanol + 1-butanol	0.33	2.42	75.00	24.89	0.05	772.5	1.74
2 allyl alcohol + 2,2,3-trimethyl, 3-pentanol	0.84	2.89	75.00	26.41	0.05	886.5	1.77
3 methanol + 2,2-dimethyl-1-butanol	0.55	2.60	75.00	26.33	0.05	805.2	1.78
4 methanol + 2-methyl-2-heptanol	0.68	2.85	75.00	26.96	0.06	1096.2	1.79
5 allyl alcohol + 2-methyl-2-heptanol	0.86	2.61	75.00	26.54	0.05	663.0	1.81
6 ethanol +2,2-dimethyl-1-butanol	0.71	2.40	75.00	25.27	0.05	630.9	2.30
7 ethanol + 2,2,3-Ttimethyl, 3-pentanol	0.79	2.90	75.00	25.34	0.05	1083.8	2.31
8 ethanol + 2-methyl-2-heptanol	0.82	2.52	75.00	25.47	0.05	755.2	2.40
9 methanol + 1-propanol, 2-methyl-	0.35	2.98	75.00	25.13	0.06	527.9	2.44
10 allyl alcohol + 2,2-dimethyl-1-butanol	0.77	2.51	75.00	26.26	0.05	597.7	2.75

Table 6.34. Liquid-liquid stability information for the mixtures of Table 6.32.

n°	mixtures	x_1	phase split
1	methanol + 1-butanol	0.33	stable
2	2,2,3-trimethyl, 3-pentanol + allyl alcohol	0.16	stable
3	methanol + 2,2-dimethyl-1-butanol	0.55	stable
4	methanol + 2-methyl-2-heptanol	0.70	stable
5	2-methyl-2-heptanol + allyl alcohol	0.14	stable
6	ethanol +2,2-dimethyl-1-butanol	0.71	stable
7	ethanol + 2,2,3-Ttimethyl, 3-pentanol	0.79	stable
8	ethanol + 2-methyl-2-heptanol	0.82	stable
9	methanol + 1-propanol, 2-methyl-	0.35	stable
10	2,2-dimethyl-1-butanol + allyl alcohol	0.23	stable

Table 6.35. Results from the verification step for the alcohol-based insect repellent case study.

n°	x_1	HB	*v*-lin	*v*-rig	R^2	RD(%)
1	0.33	PAS-PAS	2.42	2.31	0.01	4.48
2	0.16	PAS-PAS	2.89	2.77	0.02	4.52
3	0.55	PAS-PAS	2.60	2.43	0.03	6.75
4	0.70	PAS-PAS	2.85	2.60	0.06	9.37
5	0.14	PAS-PAS	2.61	2.53	0.01	3.13
6	0.71	PAS-PAS	2.40	2.31	0.01	4.27
7	0.79	PAS-PAS	2.90	2.61	0.08	11.08
8	0.82	PAS-PAS	2.52	2.33	0.04	8.29
9	0.35	PAS-PAS	2.98	2.80	0.03	6.45
10	0.23	PAS-PAS	2.51	2.49	0.00	0.77
		SD/*AAD*(%)			0.19	10.14

The last sub-task (S1-D3.5) involves optimization. The selected performance index PI was the cost. The cheapest mixture is methanol + 1-butanol.

Solution through the virtual PPD-lab

The 'Solvent mixture design' task (which corresponds to task S1-D3 in the methodology shown in Fig 5.1) was accessed. From the solvent database selection dialogue box (Fig. 6.18), both the alcohol databases (water soluble, water insoluble) were selected, as in sub-task S1-D3.1. The modelling choices were then displayed as in the dialogue box of Fig. 6.19, and they correspond to the modelling choices performed in all the case studies previously discussed in this chapter (Table 6.2). The MIXD program could be launched from the dialogue box of Fig. 6.20 (representing sub-task S1-D3.3).

The results from the MIXD program were displayed in the output file as shown in Fig. 6.20. They correspond to the mixtures of Tables 6.33-6.34.

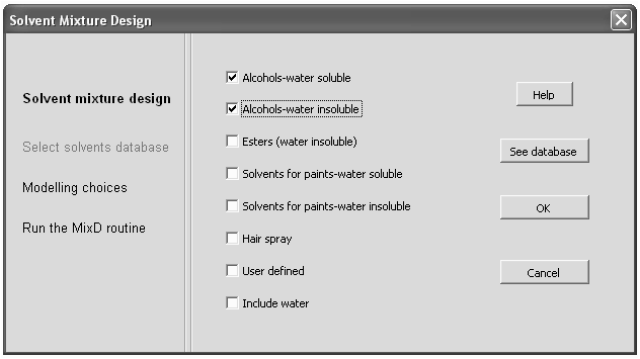


Figure 6.18. The selection of the solvent databases for the MIXD program (as in sub-task S1-D3.1).

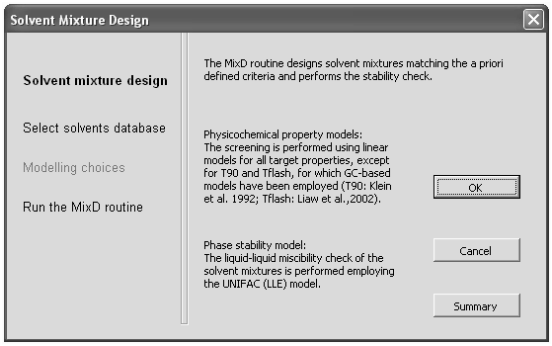


Figure 6.19. The modelling choices for the MIXD program (sub-task S1-D3.2).

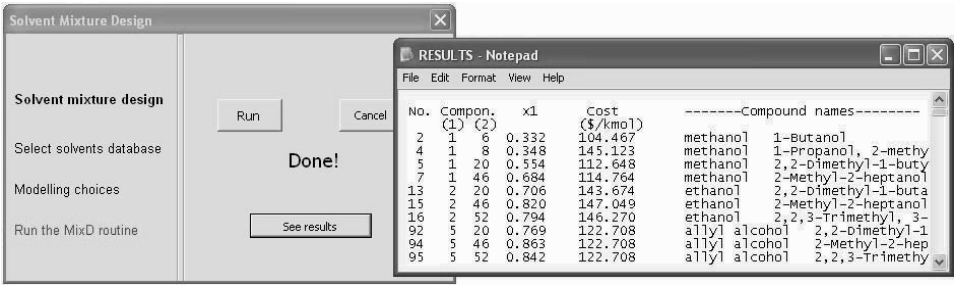


Figure 6.20. The MIXD program is launched and results are displayed (as in sub-task S1-D3.3).

The verification step has not yet been added to the virtual PPD-lab and so it needs to be performed separately. The optimization step simply involves the analysis of the results shown in the output file (Fig. 6.20).

6.4.1.4 Task S1-D4: additives identification

The quality to enhance is the scent (sub-task S1-D4.1). α/β-santalol (rarely found separately) are suitable candidates (sub-task S1-D4.2) since they are alcohols (therefore alcohol soluble) and their combined solubility parameter is 21.72 $Mpa^{½}$ (sub-task S1-D4.3) and they mtch the constraint on δ_{add} (Table 6.32). In addition, they are characterized by a green, woody scent that results to be suitable for an insect repellent lotion. They are therefore chosen as the additive for this alcohol-based insect repellent lotion (sub-task S1-D4.4).

Solution through the virtual PPD-lab

The 'Additives selection' task (which corresponds to task S1-D4 in the methodology shown in Fig 5.1) was accessed. The quality to enhance was selected in the dialogue box of Fig. 6.21 (sub-task S1-D4.1). The aroma database search dialogue box (Fig.6.22) was therefore accessed. The search for an aroma compound with a specific smell class (green, fruit, sweet,…) and with a specific common solvent was performed. In this case, the common solvent must have been an alcohol since the solvent mixture is constituted of alcohols. The smell class was selected as green (a list of smell classes is given in the 'Help' of the dialog box of Fig. 6.22). The search results are displayed in Fig. 6.22, where the aroma α/β-santalol was selected (sub-task, S1-D4.2). The properties of the compounds were collected, too, as shown in the dialogue box of Fig. 6.23 (sub-task S1-D4.3).

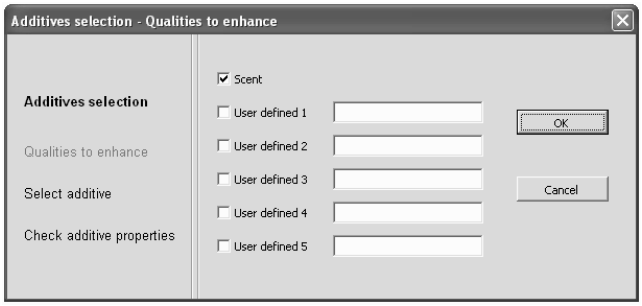


Figure 6.21. The choice of the qualities to enhance (as in sub-task S1-D4.1).

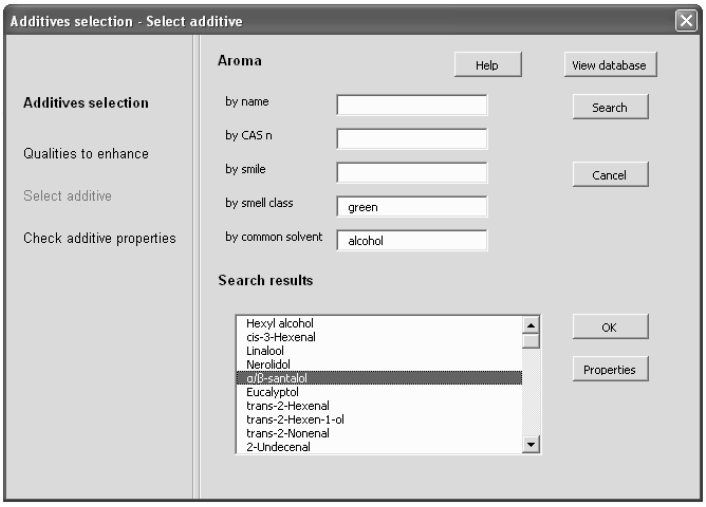


Figure 6.22. The search of the suitable aroma candidates (as in sub-task S1-D4.2).

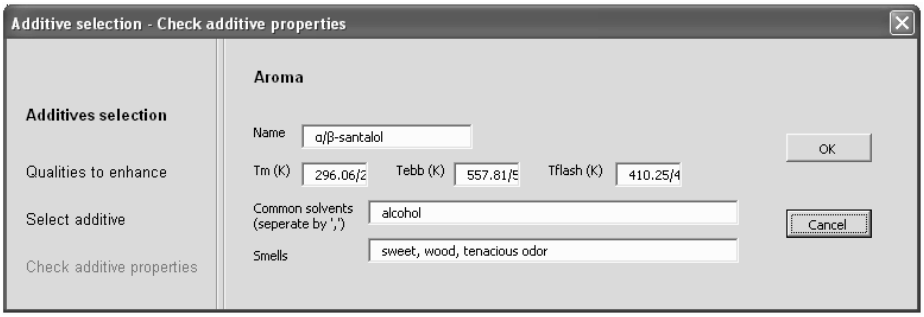


Figure 6.23. The properties of the candidate aroma compounds (as in sub-task S1-D4.3).

After α/β-santalol was selected from the dialogue box of Fig. 6.22 (sub-task S1-D4.4), and the 'Summary worksheet' was displayed, where all the information corresponding to the steps performed during the product design are shown, together with the product recipe.

6.5 Final remarks

Two main conclusions can be drawn from this chapter:

- The computer-aided stage of the methodology (consisting of collection of methods and tools, databases library, knowledge base and model library), and the framework used to develop the virtual PPD-lab are powerful instruments though which numerous alternative products can be designed, screened and/or verified.

- The experimental validation stage (S3) of the developed methodology is extremely important as it also helps to amend and refine the base case product resulting from the computer-aided stage. Experimental validation is of major importance since some of the target properties for the consumer oriented chemicals based products are not suitable for mathematical modelling. For example, the sensorial factors and the cosmetic properties such as odour, appearance, skin feeling, moisturizing effect, are very difficult if not impossible to model. In addition, sensorial factors and cosmetic properties are responsible for the success of most products on the market. Also the shelf life cannot be modelled, and it also constitutes a point of major attention for the consumers.

Some suggestions for improvements of the computer-aided stage of the methodology are discussed in this section. In §6.5.1 the possibility of improving the results of stage 1 by adding to the constraint on the Hildebrand solubility parameter, three constraints on the Hansen solubility parameters, is suggested. In §6.5.2 and 6.5.3 the possibility of including more constraints in the computer-aided screening is proposed: these constraints are on the product stability at temperatures different than the design one (§6.5.2), and on the product flammability (§6.5.3). Finally, in 6.5.4, a relation problem-cause-amendment for the type of products considered in this work is also suggested.

6.5.1 Solubility issues

In this work the mutual solubility between chemicals has been controlled setting a constraint on the Hildebrand solubility parameter. This constraint has been shown to work quite well, since through experimental validation it has been tested that most of the AIs were actually soluble in the solvent mixtures/pure solvent. Only one AI (β-carotene), in the sunscreen lotion was found to be insoluble in the designed solvent mixture/pure solvent.

Solubility investigations have been carried on the solubility of β-carotene using the Hansen solubility parameters (Hansen, 2007). Table 6.36 shows the Hansen solubility parameter values (δ_D, δ_P, δ_H, dispersive, polar and hydrogen-bonding contribution, respectively) for all the ingredients of the sunscreen lotion (except inorganic pigments, which are dispersed not dissolved). The values not available in literature have been calculated through a M&G GC-based method (Modarresi *et al.*, 2009). It has to be underlined that the accuracy of this method might be poor if the melting temperature of the chemical is far above 298 K, and this is the case of β-carotene.

Table 6.36. Hansen solubility parameters for the ingredients of the sunscreen lotion.

AI	T_m	δ_D	δ_P	δ_H
	K	$MPa^{½}$	$MPa^{½}$	$MPa^{½}$
Butyl acetate	195	15.55	4.4	6.38
Avobenzone	354-538	21.48	9.64	6.61
Octyl salicylate	< 298	17.92*	7.41*	10.80*
β-Carotene*	451-452	39.46	4.05	2.42
Vitamin A	334-336	27.89	8.19	13.42
Octocrylene*	287	16.05	11.00	6.35
Propyl paraben	368-371	17.92	9.82	12.74
Linalool	< 253	15.45	7.24	10.06

*experimental values

The fractional values of the dispersive, polar and hydrogen-bonding contribution of the Hansen solubility parameters are calculated as in Eqs. 6.36-6.38.

$$f_D = \frac{\delta_D}{\delta_D + \delta_P + \delta_H} \tag{6.36}$$

$$f_P = \frac{\delta_P}{\delta_D + \delta_P + \delta_H} \tag{6.37}$$

$$f_H = \frac{\delta_H}{\delta_D + \delta_P + \delta_H} \tag{6.38}$$

The position of the sunscreen lotion ingredients (AIs, solvent and additive) in the ternary triangular diagram of Fig. 6.24 is determined through Eqs. 6.36-6.38.

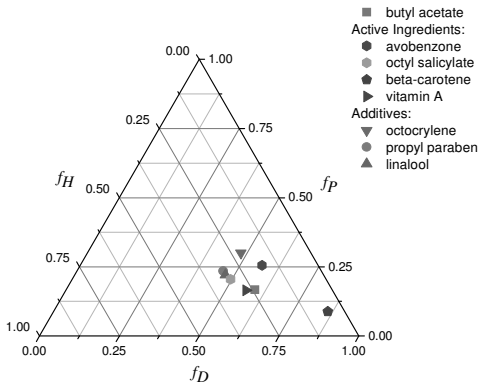


Figure 6.24. Position of the all the ingredients of the sunscreen lotion according to the Hansen solubility parameter values.

It can be noted, in Fig. 6.24, that all the ingredients are gathered in the same area, except β-carotene, which is the only ingredient that shows solubility issues. It can therefore be concluded that the screening of ingredient based on Hansen solubility

parameters results in a more effective screening than the one based on the Hildebrand parameter. In fact, the Hansen solubility parameters give many more information about the chemicals because of their three dimensional nature, which helps in comparing different kind of interactions between chemicals. Chemicals showing the same distribution of dispersive, polar and hydrogen bonding attractions are more likely to be soluble between each other.
In conclusion, in the conceptual screening performed during the computer-aided design stage of the methodology for the design and verification of formulated products, the Hansen solubility parameters should also be employed.
It has anyway to be reminded that β-carotene is not an ester while all the other AIs and additives (except linalool, the aroma compound) in the sunscreen lotion are esters. Therefore the empirical rule 'similar dissolves similar' has been shown to be successful once again.

6.5.2 Resistance to temperature changes

A consumer oriented chemicals based product like the one designed in this work should be resistant to temperature changes. Temperature changes are quite common during transportation and storage of the product.
The reaction of the product stability to changes in temperature can be predicted through solubility models. For example, predictive models such as UNIFAC can be used to predict the phase behaviour of the solvent mixture at different temperatures.
In the insect repellent case study the mixture water + isopropyl alcohol was selected. Fig. 6.25 shows the phase equilibrium isopropanol + water and the relative position of the mixture used in the final product according to UNIFAC.

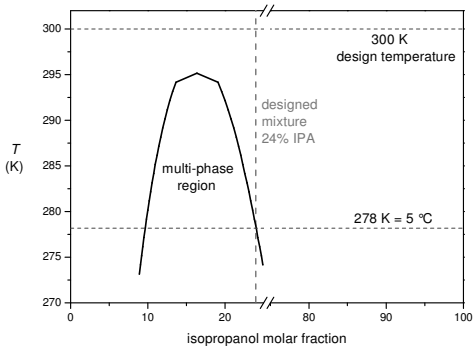


Figure 6.25. Phase equilibrium for the mixture isopropanol + water (24% molar of isopropanol) and relative position of the designed solvent mixture for the insect repellent lotion.

Fig. 6.25 reveals that the designed mixture is stable at the design temperature (300 K) since it is far away from the two phase region. But at 278 K the mixture hits the

boundaries of the unstable region. The final product (solvent mixture plus active ingredient and additives) was tested for stability at a temperature of exactly 278 K, and the test revealed the formula was still stable (§6.1.3.5). The other ingredients of the formula could have affected the phase boundaries, or the temperature should have been slightly lower than the solubility limit to cause phase split. Anyway, a preliminary study like the one in Fig. 6.25 could have been useful during the computer-aided stage of product design.
The second best solvent mixture resulting from task S1-D3 (§6.1.1.3) was ethanol + water and this mixture does not show any miscibility issues in the liquid phase for a temperature range wider than the one explored. Taking into consideration the stability at different temperatures during the computer-aided design would have driven the product developer to choose the mixture ethanol + water for the final product, which happens to be exactly the mixture employed by Bayer in Autan®.

6.5.3 Flammability issues

The flash point of the solvent mixture was not considered during the computer-aided design. This factor is extremely important since solvents are flammable chemicals and product safety is strongly affected by the solvents present in the product. The solvent mixture should have a flash point that is at least higher than the room temperature (considering that in the formulation the solvent mixture is diluted by AIs and additives that are usually not highly flammable chemicals). Models are available for the prediction of the flash point of mixtures as the model of Liaw *et al.* (2002). Using this model the flash point of the mixture 2-propanol + water (insect repellent lotion) can be predicted as function of composition, as shown in Fig. 6.26.

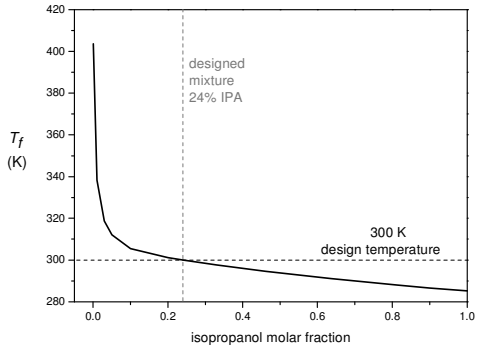


Figure 6.26. Flash point versus isopropanol composition for the mixture isopropanol + water (insect repellent lotion).

It can be noted from Fig. 6.26 that the flash point of the designed mixture is just above the safety limit of 300 K, also if water constitutes the 76% (molar) of the mixture and water is not flammable.

Once again, basing the screening on additional knowledge (data) like the flash point, a safer mixture could have been chosen during the computer-aided design stage.

6.5.4 Relation problem-cause-amendment

This joint effort of computer science and experimental validation has lead to an important achievement: the relation between problems encountered during the experimental validation and their possible cause/causes, and also suggestion on how to resolve all these problems. Table 6.37 summarizes these rules, which are applicable for the class of products taken into consideration in this work: formulations with a liquid delivery system.

Table 6.37. Relation problem-cause and suggested amendment, achieved through the case studies faced in this chapter.

problem	cause	amendment
insolubility of AI in a formula with a single AI	solvent mixture not adequate	pick another solvent mixture from the list of task S1-D3.3
insolubility of one AI in a formula with several AIs	AI not adequate	change the AI showing problems
insolubility of several AIs (multiple AIs formula)	solvent mixture not adequate	pick another solvent mixture from the list of task S1-D3.3
solubility of one or more additives	additive/additives not adequate	replace the additive/additives
the formula cannot be sprayed	viscosity and/or density are too high	try to replace the solvent mixture otherwise the product form need to be changed (cream for instance)
the product has an unpleasant colour or is turbid	one of the ingredients affects this quality factor	identify ingredients giving problems and replace it or lower its concentration
the product has an unpleasant scent	the aroma concentration is too low	increase the aroma concentration or change the aroma
	the solvent mixture has a strong smell	substitute the solvent mixture
the product is sticky or greasy	type of the ingredients and (maybe) their viscosity	identify AI or additives causing the problem and amend their concentration (if this affect the product activity, replace AI)
the product is not stable at room temperature or other temperatures	solvent mixture splits	pick another solvent mixture from task S1-D3.3
	the AI/additive is not soluble at different temperatures	
appearance/flavour/odour changes after some months	photochemical reaction	augment stabilizer concentration
	bacteria growth	augment preservatives concentration

EQUATION CHAPTER 7 SECTION 1

7

VERIFICATION CASE STUDY

In this chapter the verification option of the methodology presented in Chapter 5 (Fig. 5.1) is illustrated through a case study involving the formulation for a hair spray. The case study was proposed by Akzo Nobel (to be referred below only as 'company') and has been presented in Conte *et al.* (2010b). The main objective of the case study was to perform a preliminary verification of the product formulation and identify (design) experiments through which the final product functions, properties and phase equilibria could be verified (experiments were to be performed by the company). To carry out blind experiments on all the possible combinations between the candidate ingredients, for all composition ranges, is time consuming and expensive, and practically infeasible.

A hair spray formulation consists of a copolymer blend, solvents, propellant and additives such as neutralizing agents, plasticizers and aromas (Shah and Fernandez, 1994). The copolymer is the active ingredient that provides for the holding power. At the same time, the harsh and brittle feeling of the hair should be avoided. The solvent mixture is usually a water-based mixture of organic chemicals. Together with the propellant, the solvent mixture constitutes the delivery system. The neutralizing agents are usually alkaline and are needed to make the product water soluble (rinsable with water). Plasticizers are added to provide flexibility to the hair, and aroma compounds can also be added to enhance the cosmetic properties of the product.

In this case study, the identities of the AI and the neutralizing agent were known (given by the company), while the solvent mixture had not been defined yet, but a short-list of candidate solvents was available (from the company specifications). Because of confidentiality reasons, details such as chemical identity cannot be disclosed. However, sufficient explanations are provided for the reader to understand the main concepts. The AI is the copolymer $M1_{w1}M2_{w2}M3_{w3}$ where $M1$, $M2$ and $M3$ are three polymer repeat units and w_1, w_2 and w_3 are the weight fractions of each repeat unit, with the following values:

$$w_1 = 75.5\% \qquad w_2 = 10\% \qquad w_3 = 14.5\%$$

Note that $w_1 > > w_3 > w_2$. All the repeat units are oxygenated molecules.

The repeat unit *M*3 is actually a mixture of isomers, where the three attached groups *R*1, *R*2 and *R*3 have the distribution shown in Table 7.1. The groups *R*1, *R*2 and *R*3 listed in Table 7.1 are alkyl groups (methyl, ethyl, propyl group,…).

Table 7.1. Isomer distribution for the repeat unit *M*3 of the copolymer that constitutes the AI. '–CH3' is the methyl group, while '> CH3' represents alkyl groups with more carbon atoms than the methyl group (i. e., ethyl, propyl,…)

***R*1**	***R*2**	***R*3**	**Distribution (%)**
$-CH_3$	$-CH_3$	$> CH_3$	31
$-CH_3$	$> CH_3$	$> CH_3$	67
$> CH_3$	$> CH_3$	$> CH_3$	2

The copolymer concentration (w_{Cop}) in the system is fixed to 5% by weight, which is the concentration that ensures the desired functions for the hair spray (curl retention, shine, ...). A range of values for the copolymer molecular weight (*Mw*) and for the polydispersity index (*PDI*) were also given:

$$55000 \le Mw \le 75000; \quad 3 \le PDI \le 3.5$$

The polidispersity index is a parameter that lumps the variance of the molecular weight distribution curve of a polymer. As the value of *PDI* increases, the *Mwd*-curve becomes wider making the amount of polymer with a short chain length larger.
The list of solvents includes 5 candidates: *A*, *B*, *C*, *D* and *E*. The additive is a neutralizing agent (to make the product water rinsable).
The goals of this case study were:

- To define the problem: given the performance criteria (the consumer needs), translate the needs into physicochemical properties and set the constraints;
- To build an appropriate suite of property models for the calculation of the phase behaviour of the copolymer in solvents;
- To investigate the phase equilibria of the system: phase boundaries for the AI in single solvents and solvent mixtures. In addition, the phase behaviour dependence on the copolymer molecular weight and molecular weight distribution needs to be analyzed in order to identify the critical values (leading to the largest immiscibility gap);
- To identify a pure solvent or a single liquid phase binary mixture of solvents that satisfies the performance criteria set by the market/regulatory. At the same time, the pure solvent/solvent mixture has to dissolve the AI. Only if a pure solvent or a binary solvent mixture did not match the required targets, ternary (quaternary,…) solvent mixtures were to be considered;

- Finally, to verify that the given neutralizing agent is compatible with the best performing blend copolymer-solvents previously identified and that it can make the copolymer water rinsable.

The temperature range of interest for the case study is around 300 K, the room temperature, which represents also the temperature at which the product has to be employed. Fig. 7.1 highlights the tasks of the overall methodology for formulation design and verification that are involved in this cases study.

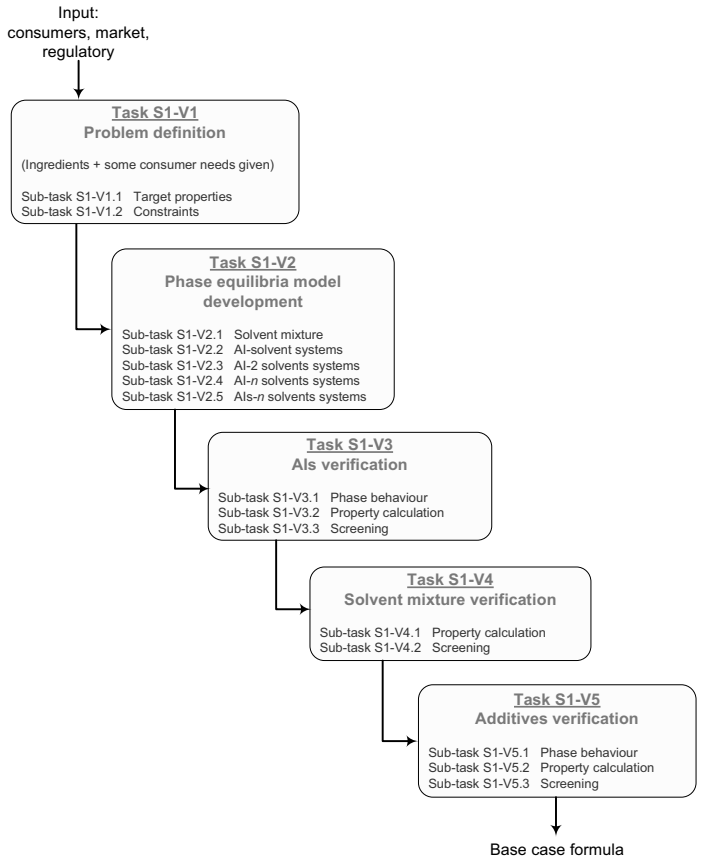


Figure 7.1. The tasks of the overall methodology for formulation design and verification (Fig. 5.1, Chapter 5) that are involved in this cases study.

7.1 Task S1-V1: problem definition

A list of ingredients was given by the company, together with the performance criteria.

Market surveys indicate that consumers want a hair spray that gives a good curl retention and holding power, without giving a harsh and brittle feeling to the hair

(Varco and Williams, 1986). Shine and lustre of the hair are also required attributes (Shernov, 1991). It must be possible to easily rinse the product with water (Varco and Williams, 1986) and products with short drying time are always preferred. In addition, the spray does not have to retain electric charge, which is responsible of the 'electric look' (Shernov, 1991). Flammability and toxicity concerns must also be taken into considerations (Shernov, 1991). Environmental friendly products are preferred therefore VOCs (Volitile Organic Compounds) emissions must be below the regulated limit (Shah and Fernandez, 1994; Shernov, 1991).
The copolymer (AI) should be uniformly dissolved in the solvent mixture, to ensure the product effectiveness. In fact, the function of retaining the hair-curls (for which the polymer is responsible) cannot be guaranteed if the polymer deposits on the bottom of the product dispenser.
The curl retention/holding power, shine and lustre are provided by the AI. The tacky or gummy feeling also depends on the AI. The solvent mixture is responsible for the product drying time, flammability, toxicity and VOCs emissions. The static charge of the solvent mixture should be kept low. Actually the polymer affects the static charge of the overall formulation, but the copolymer identity and/or structure are not under evaluation.

7.1.1 Sub-task S1-V1.1: target properties

The drying time of the solvent mixture is related to the parameter T_{90}, which is the time required by the 90% by weight of the solvent mixture to evaporate. The toxicity can be related to the parameter LC_{50}, which is the lethal concentration of a pure chemical or mixture that causes 50% of death in a fathead minnow population. The dielectric behaviour of the product is strictly related to the dielectric constant ε. The closed-cup flash point (T_f) accounts for the product flammability. The VOCs emissions depend on the alcohol concentration (w_{OH}) of the solvent mixture (if any alcohol is present).
To ensure that the copolymer is uniformly dispersed in the solvent mixture, constraints need to be applied on the delta Gibbs energy of mixing (ΔG^{mix}) and on the tangent plane distance (*TPD*, see Chapter 4, §4.1.3).

7.1.2 Sub-task S1-V1.2 constraints

Using the knowledge base, constraint values were set in this sub-task. The drying time constraint was chosen considering information from patented product prototypes (Morawsky and Martino, 1997). The toxicity was fixed to low values since a hair spray is a cosmetic product applied directly on the body (the higher the value of LC_{50}, the least toxic the compound). The dielectric constant constraint was fixed close to the value of water to avoid the dielectric behaviour. The flash point has to be at least higher than the room temperature for safety reasons; hence the lower bound was fixed at 300 K. In order to ensure the liquid phase stability of the product (solubility of the

AI in the solvent mixture and the liquid-liquid miscibility of the solvent constituting the solvent mixture), the same constraints applied for the case studies of Chapter 6 are employed: the function $\Delta G^{mix}/RT$ should be negative and the tangent plane condition should be satisfied, at the specific product composition. Differently from the case studies in Chapter 6, here these two constraints were applied on the overall product formula, and not only on the solvent mixture.
The overall mixture should contain less than 80% (by weight) of alcohol in order to limit the VOCs emissions.
Numerical values for the constraints were set as in Eqs. 7.1-7.7. Eq. 7.6-7.7 are the two conditions for single liquid phase.

$$480 \le T_{90} \le 960 \quad (\mathrm{s}) \tag{7.1}$$

$$LC_{50} \ge 0.1 \quad (\mathrm{mol/l}) \tag{7.2}$$

$$50 \le \varepsilon \le 70 \tag{7.3}$$

$$T_f \ge 300 \quad (\mathrm{K}) \tag{7.4}$$

$$w_{OH} \le 80\% \tag{7.5}$$

$$\frac{\Delta G^{mix}}{RT} < 0 \tag{7.6}$$

$$TPD \ge 0 \tag{7.7}$$

Table 7.2 lists the performance criteria, the corresponding target properties and the constraints values.

Table 7.2. Target property constraints for the hair spray case study. LB and UB are the lower and the upper bound, respectively. UoM stands for Unit of Measure.

target property	symbol	UoM	LB	UB
evaporation time	T_{90}	s	480	960
lethal concentration	LC_{50}	mol/l	0.1	$+\infty$
dielectric constant	ε	-	50	70
flash point	T_f	K	300	$+\infty$
alcohol concentration	w_{OH}	kg/kg	0.0	0.8
Δ Gibbs energy of mixing	$\Delta G^{mix}/RT$	-	$-\infty$	0
tangent plane distance	TPD	-	0	$+\infty$

7.2 Task S1-V2: phase equilibria model development

For the problem defined above (task S1-V1), the appropriate property models were developed as part of task S1-V2.

The modelling problem was decomposed according to the guidelines given in §5.1.3.2. Fig. 7.2 illustrates the decomposition for this particular case study, where the AI is a copolymer constituted of 3 repeat units. At first, the solvent mixture was analyzed (sub-task S1-V2.1). At sub-task S1-V2.2, the systems constituted of the polymer *P*1 (composed by the repetition of the monomer *M*1) in single solvents were considered. This was repeated for polymers *P*2 and *P*3 (constituted of the repetition of monomer *M*1 and *M*2, respectively). Then, the systems formed by polymer P_i and solvent binary mixtures were analyzed (sub-task S1-V2.3). Since the problem only required the identification of a pure solvent or the design of a binary solvent mixture, sub-task S1-V2.4 (behaviour of each polymer in multicomponent mixtures) was not performed. Finally in sub-task S1-V2.5, the three repeat units *M*1, *M*2 and *M*3 were combined to form the copolymer $M1_{w1}M2_{w2}M3_{w3}$ and the phase behaviour in single solvents and solvent binary mixtures was considered. The dashed arrows in Fig. 7.2 highlight the fact that after performing task S1-V2.2, the copolymer solubility in pure solvents can be simulated, or that after task S1-V2.3 the copolymer phase behaviour in binary mixtures can be calculated (as in this case study).
Average values for the molecular weight and polydispersity index needed to be selected in order to perform the calculations required by the phase equilibria task. The following values were selected: $Mw = 65000$; $PDI = 3.5$.

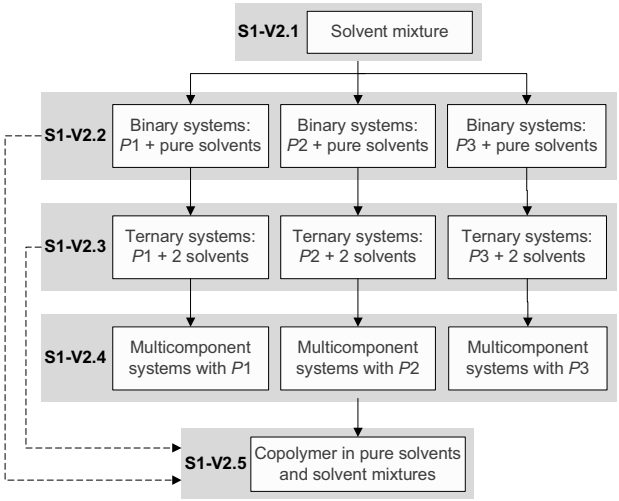


Figure 7.2. The phase equilibria task S1-V2 applied to the hair spray case study.

Note that in task S1-V3 (§7.3) a sensitivity analysis was to be performed to analyze the effect of the *Mw* and the *PDI* on the phase behaviour. With this analysis, it was be possible to identify the conditions at which the two phase regions for the copolymer-solvents systems have their maximum extension. In fact, to ensure the solubility of all polymer chains (that have different length since $PDI \neq 1$) the worst case scenario (the largest immiscibility gap) needed to be considered.

7.2.1 Sub-task S1-V2.1: solvent mixture

Solvent mixtures were analyzed here. Since the objective was to identify a pure solvent or to design a solvent mixture that matched the *a priori* defined criteria and dissolved the AI, only pure solvents and binary solvent mixtures were analyzed in this sub-task. Only if no pure solvent or binary mixtures of solvents did not satisfy the above criteria, multicomponent solvent mixtures (ternary, quaternary,…) would have been considered.

Ten solvent systems were analyzed in this sub-task. These are all the possible binary combinations between the five solvents in the shortlist (*A*, *B*, *C*, *D*, *E*). At first, a search for experimental data was performed to find information on liquid-liquid miscibility. When experimental LLE/solubility data were not available, models (in this case study, NRTL, UNIQUAC or UNIFAC) were employed to estimate the liquid-liquid miscibility.

Liquid-liquid equilibrium data were found only for the system (*A*-*E*) (Sørensen and Arlt, 1979). For the other systems, models were employed to identify any miscibility issue. LLE was observed for the following binary mixtures: (*A*-*D*), (*C*-*D*) and (*D*-*E*). Fig. 7.3 shows the LLE phase diagrams for the four solvent systems having immiscibility regions.

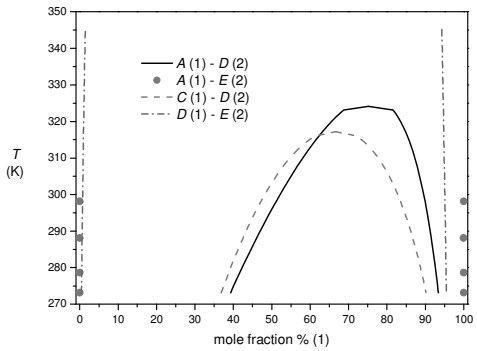


Figure 7.3. LLE for the four binary mixtures of solvents having immiscibility regions.

7.2.2 Sub-task S1-V2.2: polymer in pure solvent

The total number of systems to analyze in this step was 15, since the phase behaviour of each polymer (constituted of the repetition of the monomers *M*1 *M*2 and *M*3) in single solvents needed to be considered. Experimental data for the systems under consideration are very scarce. For this reason it was decided to employ a predictive model for the calculation of the phase behaviour. The GC-Flory Equation of State (Holten-Andersen *et al.*, 1987; Chen *et al.*, 1990; Bogdanic and Fredenslund, 1994) was selected. This model had been presented in §3.6.

7.2.2.1 The GC-Flory EoS

The accuracy of the GC-Flory EoS in the prediction of VLE is quite good (Bogdanic and Fredenslund, 1994) and qualitatively correct. LLE behaviour can be obtained (Saraiva, 1995), but the model has not been sufficiently tested on systems containing oxygenated polymers. In fact Saraiva (1995) took into consideration only the system polyethylene glycol (PEG) + water. Before proceeding with the calculations related to the cases study under consideration, the performance of the GC-Flory EoS for the prediction of LLE for oxygenated polymers was evaluated.

Experimental data for four systems containing Poly(n-butylmethacrylate) (PnBMA, *Mw* = 11600) in solvents (ethanol, methanol, pentane and octane) were found (Saraiva *et al.*, 1995). The GC-Flory EoS was employed for the prediction of LLE for these four polymer-solvent systems, and results are compared with the experimental data (Fig. 7.4).

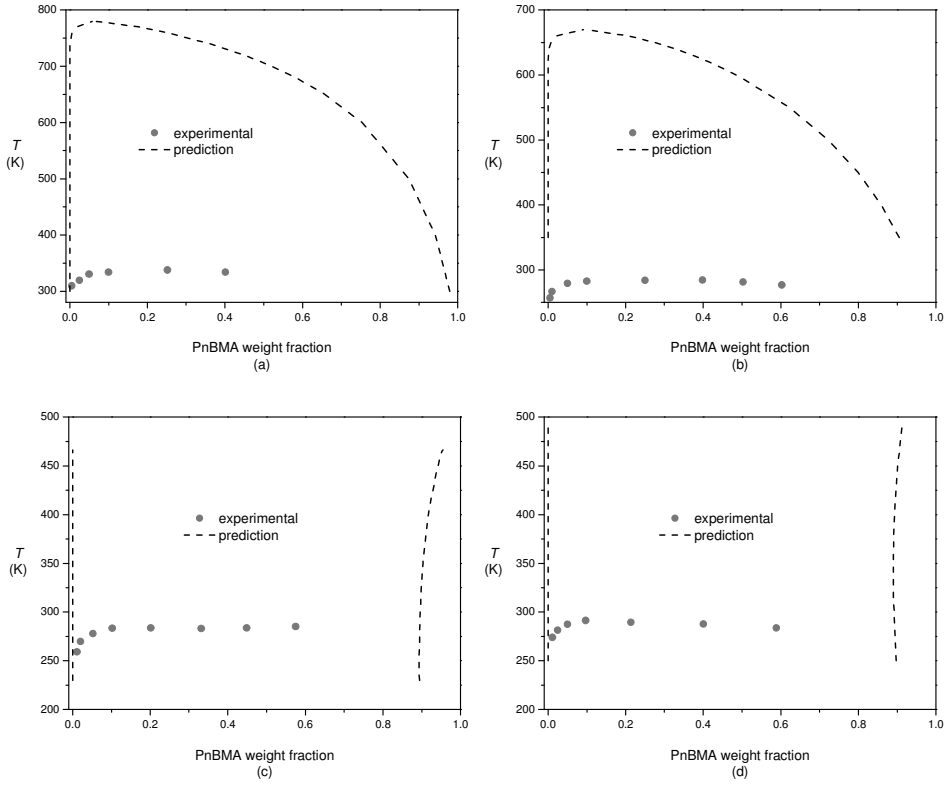


Figure 7.4. LLE phase diagrams for the systems: (a) PnBMA-methanol; (b) PnBMA-ethanol; (c) PnBMA-pentane; (d) PnBMA-octane.

It can be noted that the GC-Flory EoS was able to predict the UCST (Upper Critical Solution Temperature) behaviour for PnBMA in methanol and ethanol, but with a very large deviation from the experimental data. For the systems PnBMA in hydrocarbons (pentane and octane) the GC-Flory EoS predicts the hour glass behaviour instead of the experimental UCST diagrams. Therefore, it was concluded that this model is not able to predict LLE in systems containing oxygenated polymers, either qualitatively or quantitatively. Another model, able to handle oxygenated polymers was therefore necessary. The FV-UNIQUAC (Bogdanic and Vidal, 2000; Bogdanic, 2001) was selected since it is able to correlate LLE for polymer solutions with satisfactory accuracy, including oxygenated polymers.

7.2.2.2 The FV-UNIQUAC model

The FV-UNIQUAC model that was presented in §3.7 (Chapter 3) was considered here. This model is based on segmental interactions, therefore, it is predictive only if the necessary segment interaction parameters are available. The model requires, as input, the densities of the compounds involved, the van der Waals volumes, the molecular surface parameters and the segmental interaction parameters. For the volume/density calculation the following model were used:

- Solvent: DIPPR databank (Daubert and Danner, 1985); if the solvent was not present in such databank, a group contribution model (GCVOL method) was employed to calculate the molar volume (Elbro *et al.*, 1991; Ihmels and Gmehling, 2003);
- Polymer: Polymer Solution Handbook for the Tait equation parameters (Danner and High, 1992); if data were not available for the polymer under consideration, the GCVOL method was employed (it is applicable to solvents, oligomers and also polymers).

The Van der Waals volumes and the molecular surface parameters were calculated with the group contribution model of Bondi (1968).
When the needed segmental interaction parameters were not available, they were regressed on experimental data. When experimental LLE data were regressed, the objective function minimized in the optimization algorithm was:

$$objF = \sum_{k}^{NE} \sum_{i}^{NC} \left[\left(a_i^I - a_i^{II} \right)^k \right]^2 \tag{7.8}$$

where a_i^I and a_i^{II} are the calculated activities of compound i in phase I and in phase II, in the k-experiment. NE is the number of experimental points (a_i^I-a_i^{II}) and NC the number of compounds in the mixture. When experimental solubility data (x_i-a_i, x_i-$\ln(\gamma_i)$) were regressed, the objective function became:

$$objF = \sum_{k}^{NE} \sum_{i}^{NC} \left[\left(a_i^{exp} - a_i^{est} \right)^k \right]^2 \tag{7.9}$$

where a_i^{exp} and a_i^{est} are the experimental and the calculated activities of compound *i*, respectively. When experimental values of the activity coefficient were available, the difference between the natural logarithm of the experimental and calculated values was minimized. The Levenberg-Marquardt method was used for the parameter regression.

7.2.2.3 The parameter regression

None of the needed interaction parameters for the system under consideration were available in the FV-UNIQUAC parameter tables (Bogdanic and Vidal, 2000; Bogdanic, 2001), therefore all the necessary parameters had to be regressed using experimental data. Data were found for the systems (*P*1-*A*), (*P*1-*B*) and (*P*1-*C*) (Hao, Elbro and Alessi, 1992; Wibawa *et al.*, 2002). Qualitative information were also found for the systems (*P*1-*A*), (*P*1-*B*): these systems are immiscible (Danner and High, 1992).

For all the other systems, no data were found in the literature. In this case, the same hypothesis used by Bogdanic and Vidal (2000) was employed: "the energetic interactions between segments are not significantly dependent on molar mass". According to this hypothesis, experimental data at the monomer level could be employed for the parameter regression when data for polymer-solvent systems were not available. This approach is further explained below through Fig. 7.5.

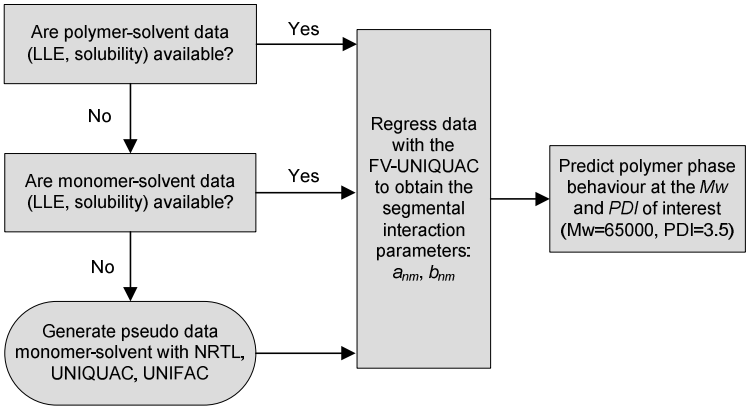


Figure 7.5. The approach employed in this case study for the regression of the interaction parameters. The approach is based on the hypothesis of Bogdanic and Vidal (2000).

According to Fig. 7.5, at first literature is searched to find LLE or solubility data for polymer-solvent systems. The data can be for a polymer at a different molecular

weight or polydispersity than the ones desired. If data are available, interaction parameters are regressed and the phase behaviour of the polymer with the desired *Mw* and *PDI* is predicted. If data at the polymer scale are not available, monomer-solvent data are employed for the parameter estimation. If not even data at low molecular weight scale are available in the literature, pseudo-data monomer-solvent are generated with well known excess Gibbs energy models that can describe the systems under evaluation. In this work, the UNIFAC GC-based method (Magnussen *et al.*, 1981) was employed for the pseudo-data generation, since other models (NRTL, UNIQUAC) did not have the required interaction parameters.

The approach highlighted in Fig. 7.5 was tested, to verify the assumptions used.

The approach was tested on the system (*P*1-*A*) since for this system two sets of data were found: one data set for the polymer-solvent system at 313.15 K, and another set at low molecular scale (*M*1-*A*) in the range of temperature 293.15 – 328.4 K. The first data set (polymer level) could not be used for regression since the temperature is different from the one of interest (300 K). It could, instead, be employed for verification, after the interaction parameters had been obtained through regression of the experimental data for the system (*M*1-*A*). Fig. 7.6a shows the regression results for the system (*M*1-*A*) while Fig. 7.6b shows the predicted activity of solvent *A* in polymer *P*1 at 313.15 K, obtained by employing the parameters regressed using employing the monomer-solvent system data. In the same plot (Fig. 7.6b), the experimental data are also shown.

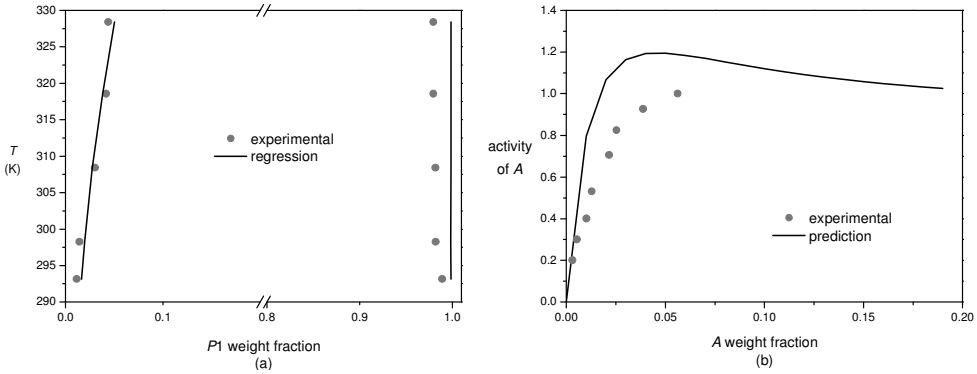


Figure 7.6. (a) Regression of experimental data (*M*1-*A*), from which the segmental interaction parameters were obtained. (b) Prediction of the activity of solvent *A* in the polymer *P*1 and comparison with experimental data at 313.15 K.

The agreement is satisfactory, considering that the phase behaviour of the polymer-solvent system was extrapolated from low molecular weight data. Therefore the approach of Fig. 7.5 was adopted for the systems for which no polymer-solvent data could be found.

In this work temperature dependent data were usually regressed, in order to obtain all the four segmental interaction parameters (a_{12}, a_{21}, b_{12}, b_{21}) for all the systems. In case the model showed difficulties in fitting the experimental data, only data at 300.15 K were regressed. Table 7.3 summarizes the information about the polymer-solvent systems under consideration. Information about the problems encountered during the regression for some of the systems under consideration ((*P*1-*E*), (*P*3-*A*)) are also given, in the last column. These problems are discussed in §7.2.2.4.

Table 7.3. Type of experimental and pseudo-experimental data for polymer-solvent systems, temperature dependence of the data (iso-*T*: isothermal data, 300 K; *T*-dep: temperature dependent data) and sum of square residuals (value of the objective function in Eq. 7.8 or Eq. 7.9). Information about problems encountered during the parameter regression are shown in the last column.

nº	(1)	(2)	Experimental data	*T*-dependence	ΣR^2	Problems
1	*P*1	*A*	LLE (*M*1-*A*); w_i-a_i (*P*1-*A*)	*T*-dep	8.57	-
2	*P*1	*B*	w_i-a_i (*P*1-*B*)	*T*-dep	5.18E-2	-
3	*P*1	*C*	w_i-a_i (*P*1-*C*)	*T*-dep	1.04E-1	-
4	*P*1	*D*	pseudo x_i-ln(γ_i) (*M*1-*D*)	*T*-dep	7.34E-3	-
5	*P*1	*E*	pseudo x_i-ln(γ_i) (*M*1-*E*)	*T*-dep	1.26E-1	yes
6	*P*2	*A*	pseudo x_i-ln(γ_i) (*M*2-*A*)	*T*-dep	7.77E-1	-
7	*P*2	*B*	pseudo x_i-ln(γ_i) (*M*2-*B*)	*T*-dep	1.28E-5	-
8	*P*2	*C*	pseudo x_i-ln(γ_i) (*M*2-*C*)	*T*-dep	3.58E-4	-
9	*P*2	*D*	pseudo x_i-ln(γ_i) (*M*2-*D*)	*T*-dep	6.96E-5	-
10	*P*2	*E*	pseudo x_i-ln(γ_i) (*M*2-*E*)	*T*-dep	1.40E-4	-
11	*P*3	*A*	pseudo x_i-ln(γ_i) (*M*3-*A*)	iso-*T*	7.11	yes
12	*P*3	*B*	pseudo x_i-ln(γ_i) (*M*3-*B*)	*T*-dep	2.5E-1	-
13	*P*3	*C*	pseudo x_i-ln(γ_i) (*M*3-*C*)	*T*-dep	2.28E-1	-
14	*P*3	*D*	pseudo x_i-ln(γ_i) (*M*3-*D*)	*T*-dep	7.54E-1	-
15	*P*3	*E*	pseudo x_i-ln(γ_i) (*M*3-*E*)	*T*-dep	2.7E-1	-

For the systems (*P*1-*B*), (*P*1-*C*) experimental solubility data were regressed and the results of the regression are highlighted in Figs. 7.7a and 7.7b, respectively. For the system (*P*1-*D*), pseudo-data involving monomer solvents (*M*1-*D*) were regressed, and the results are highlighted in Fig. 7.7c. It can be noted that the FV-UNIQUAC model is able to describe these low molecular weight systems quite accurately.

All the systems involving monomer *M*2 required the generation of pseudo-experimental data. With the exception of the system (*M*2-*A*) for which some off-set could be noticed, these systems could be very well described by the FV-UNIQUAC model (see Figs. 7.8b-e, which show the data were very well fitted).

Monomer *M*3 is a mixture of isomers (Table 7.1). No experimental solubility data for the polymer *P*3/monomer *M*3 in solvents could be found. Therefore, pseudo-experimental data were generated also for these systems. To simplify the problem, a representative molecular structure of monomer *M*3 was selected. UNIFAC could not entirely describe the molecule *M*3 with the available groups, since one group is missing in the UNIFAC-LLE parameter tables. Looking at UNIFAC-LLE parameter tables, it resulted obvious that the missing group *k* (it is a sub-group in the UNIFAC

terminology) would clearly belong to an already existing main-group (main-group 14: 'COOC').

Hence, an additional sub-group *k* was created, with the following characteristics:

- The volume and surface parameters R_k and Q_k for the new group *k* were calculated from the Van der Waals group volume and areas (V_k and A_k) given by Bondi (1968);
- The interaction parameters of the existing main-group 'COOC' with other main groups were used for this new sub-group *k* (since the sub-group *k* is clearly part of the existing main-group 'COOC').

After this addition to the UNIFAC-LLE parameter table, pseudo-experimental data could be generated. Fig. 7.9 shows the performance of the regressed parameters for systems involving monomer *M*3. The correlations are not as successful as for the systems involving monomer *M*2, mainly for (*M*3-*B*).

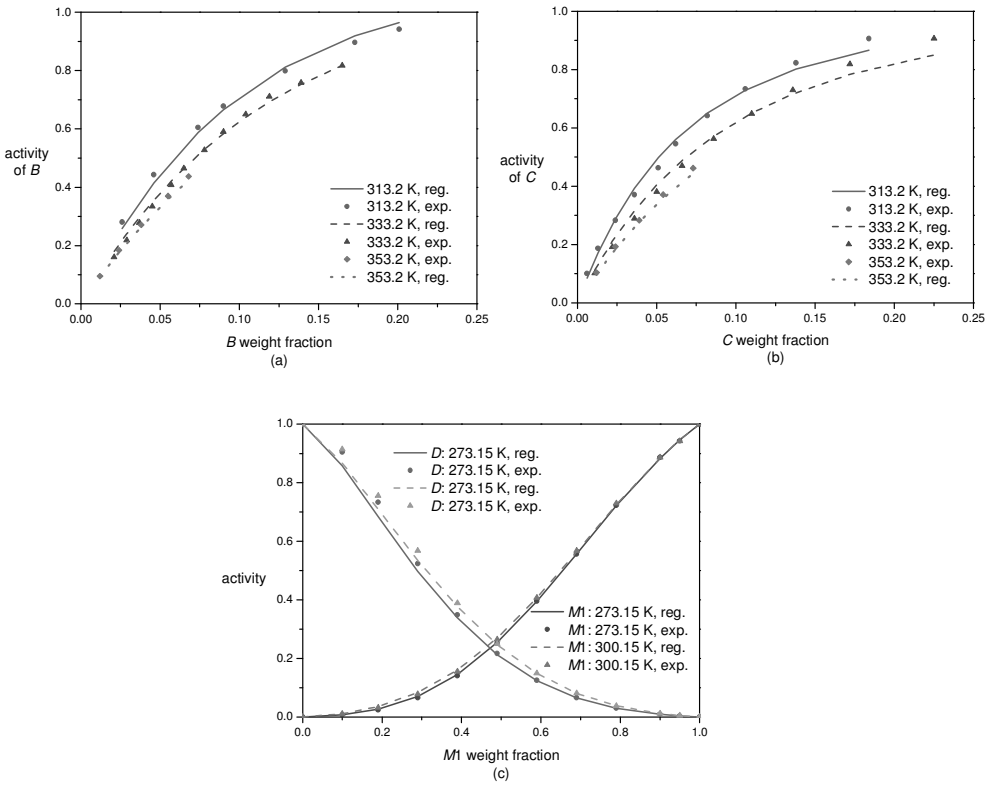


Figure 7.7 . Regression of experimental solubility data for the system (*P*1-*B*) (a), and (*P*1-*C*) (b). Regression of pseudo-experimental solubility data for the system (*M*1-*D*) (c). 'Reg.' stands for regression and 'exp.' for experimental (or pseudo-experimental).

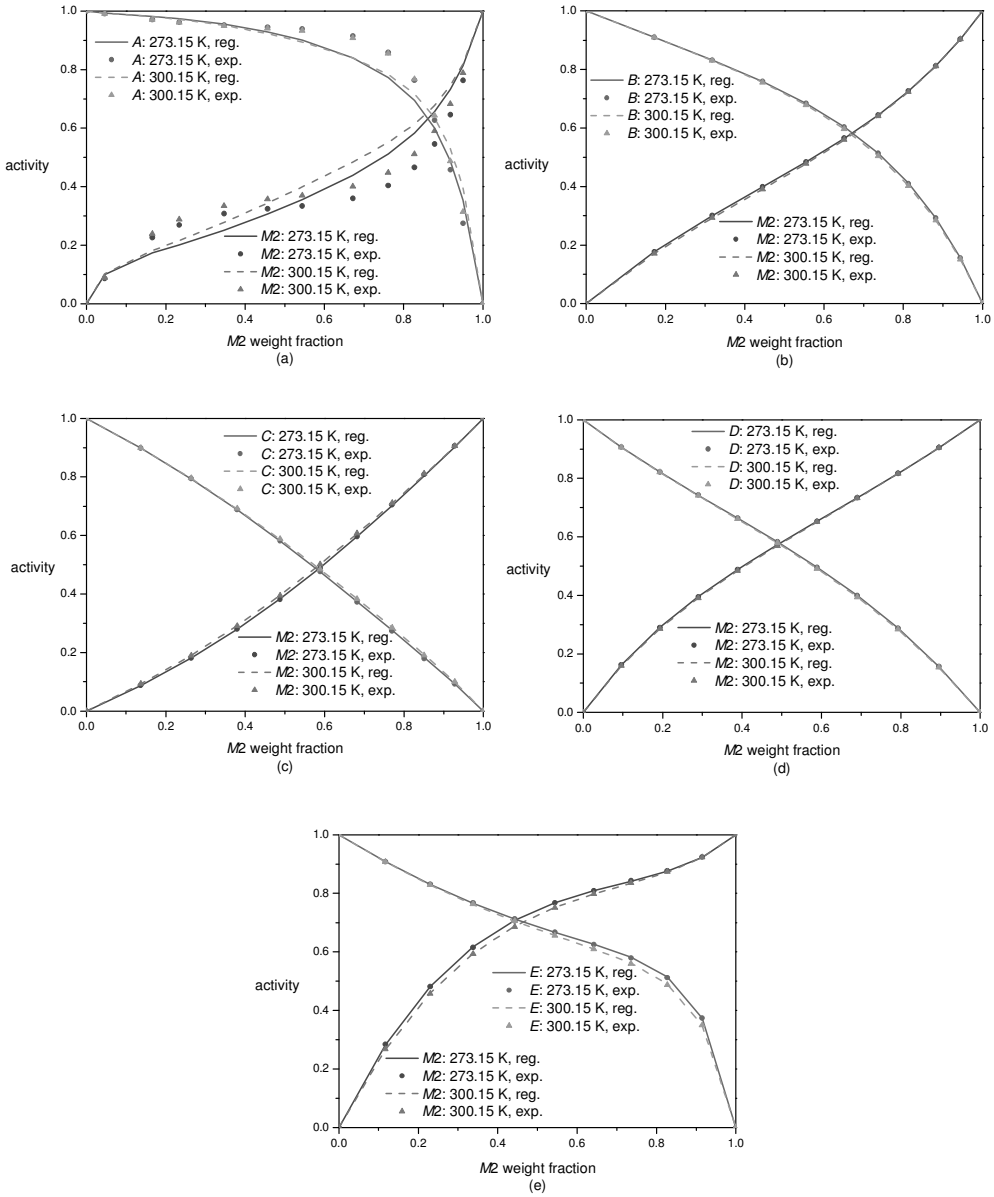


Figure 7.8. Regression of pseudo-experimental solubility data for the systems: (a) (*M*2-*A*); (b) (*M*2-*B*); (c) (*M*2-*C*); (d) (*M*2-*D*); (e) (*M*2-*E*).

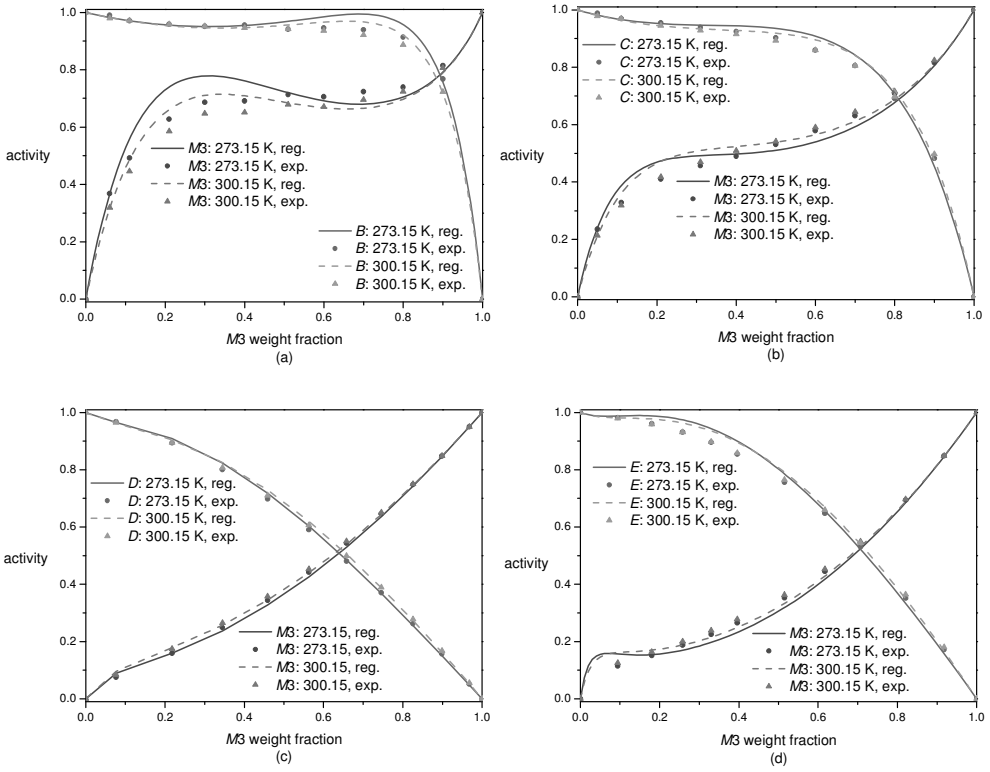


Figure 7.9. Regression of pseudo-experimental solubility data for the systems: (a) (*M*3-*B*); (b) (*M*3-*C*); (c) (*M*3-*D*); (d) (*M*3-*E*).

7.2.2.4 Modelling problems polymer-solvent

Modelling problems were encountered for the systems: (*P*1-*E*), (*P*3-*A*). Justifications and solutions to these problems are highlighted in the text below.

System P1-E

For the system (*P*1-*E*) pseudo-experimental data for the system (*M*1-*E*) (generated with the UNIFAC model) were regressed (Fig. 7.10a). An irregularity in the data and therefore also in the regression curve could be noted for low concentrations of monomer *M*1 ($w_{M1} < 0.2$). When predicting the activity coefficient for the polymer-solvent system (*P*1-*E*) employing the regressed interaction parameters, this irregularity led to very high infinite dilution activity coefficient values ($\ln(\Omega) \sim 650 \Rightarrow \Omega \sim 2.0E+282$). This caused numerical problems in the calculation of the liquid-liquid equilibrium.

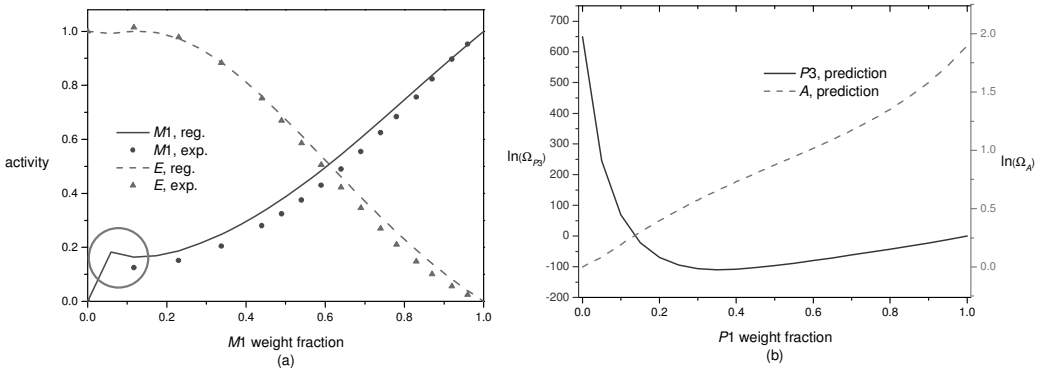


Figure 7.10. (a) Regression results for the system (*M*1-*E*) when all the pseudo-experimental data generated with UNIFAC were taken into consideration and (b) prediction of the activity coefficient trends for the system (*P*1-*E*). The red circle in (a) highlights the irregularity in the activity of *M*1. T = 300.15 K.

It has to be underlined that all the other systems of polymer *P*1 in solvents showed low values of the activity coefficients in almost all the composition range (indicated by negative values of the logarithm of the activity coefficients). This observation suggested that the trend of Fig. 7.10b was wrong.

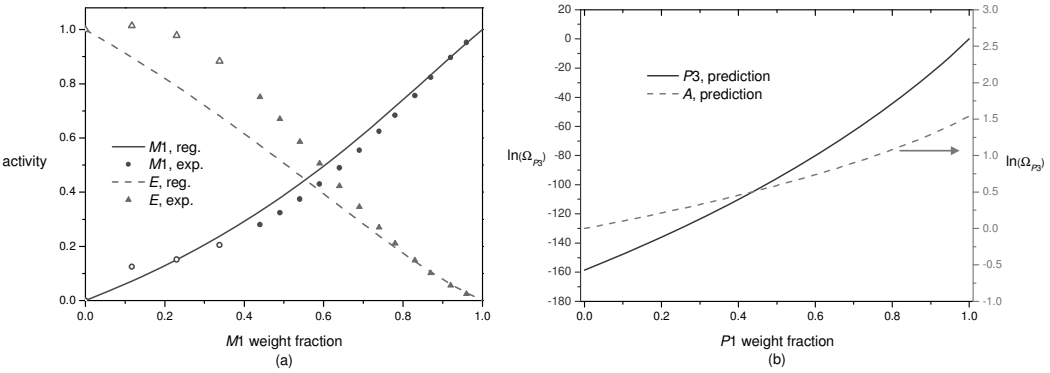


Figure 7.11. (a) Regression results for the system (*M*1-*E*) when some of the pseudo-experimental data generated with UNIFAC (the ones represented by an empty symbol) were not considered. (b) Prediction of the activity coefficient trends fro the system (*P*1-*E*). T = 300.15 K.

It was decided to perform another parameter regression on a reduced data set, in which the pseudo-experimental data corresponding to a weight fraction w_{M1} lower than 0.4 were not considered (these data points are indicated by empty symbols in Fig. 7.11a). The reduced data set was then employed for regression of the interaction parameters (Fig. 7.11a). The irregularity at low concentrations of monomer *M*1 was avoided, at the expense of having a poorer prediction of the activities at low concentrations. In addition, the trend of the polymer activity coefficient (Fig. 7.11b) became comparable to the behaviour of the other polymer-solvent systems. Temperature dependent pseudo-experimental data were regressed for this system, as reported in Table 7.3. Figs. 7.10-7.11 are an example of the regressions for only one temperature (300.15 K), in order to show the regression procedure for this particular system.

System P3-A

The FV-UNIQUAC model was found to have a limitation in fitting the pseudo-experimental data generated with UNIFAC for the system (*M*3-*A*). In fact, a poor fit of the pseudo-experimental data (isothermal data, 300.15 K) was observed (Fig. 7.12).

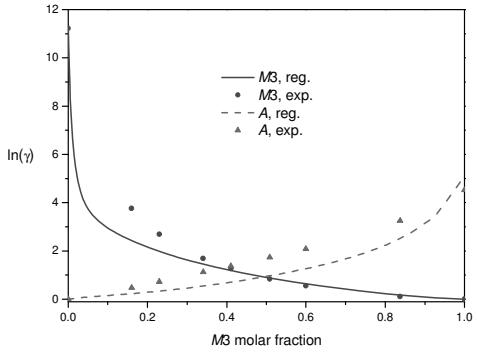


Figure 7.12. Regression of pseudo-experimental solubility data for the system (*M*3-*A*) at 300.15 K; due to the very large values of the activity coefficient for low monomer composition, it is preferred to show the activity coefficient trend versus the concentration of monomer *M*3.

The segmental interaction parameters obtained from the regression were employed to predict the activity coefficients of the system (*P*3-*A*) as shown in Fig. 7.13a. The polymer infinite dilution activity coefficient could not be calculated since the value is too high (out of the range of the computer). In addition, some numerical problems appeared for high polymer weight fraction (see Fig. 7.13b, which is the enlargement of Fig. 7.13a for w_{P3} = 0.95-1.0). In fact, the infinite dilution coefficient of solvent *A* ($\ln(\Omega_A^{\infty})$) should not drastically decrease to a value close to zero. Additionally, the

logarithm of the polymer activity coefficient ($\ln(\Omega_{P3})$) oddly decreases and then increases between polymer composition w_{P3} = 0.95-1.00.

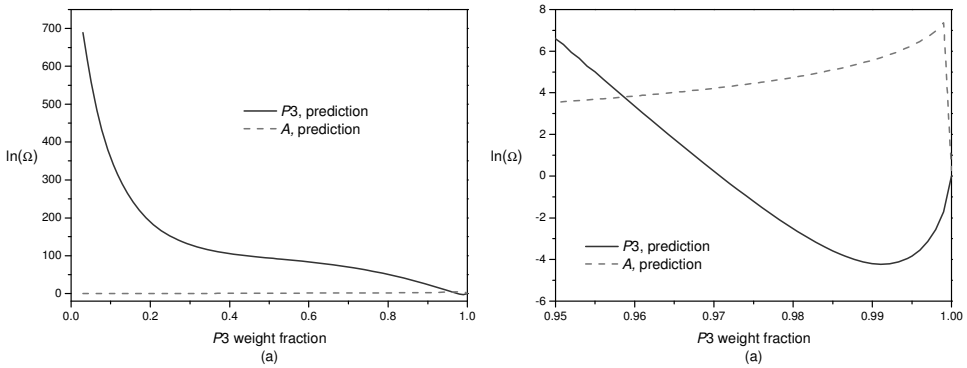


Figure 7.13. (a) Predicted activity coefficient trend for the system *P*3-*A*, and (b) enlargement of (a) for high polymer *P*3 concentrations. T = 300.15 K.

These numerical problems could be resolved by fine-tuning the interaction parameters using a modified set of data points for the system (*P*3-*A*):

- A series of points on the polymer activity coefficient trend of Fig. 7.13a was selected as a set of data points on which to fine-tune the segmental interaction parameters;
- The data points corresponding to the infinite dilution activity coefficients were extrapolated using a 6th-order polynomial trend line on the selected data;
- The regression was performed on the modified data points for the system (*P*3-*A*). Results are highlighted in Fig. 7.14.

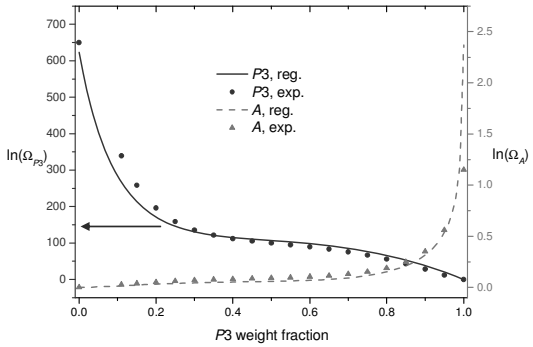


Figure 7.14. The modified trends of the activity coefficient for the system *P*3-*A* were regressed in order to fine-tune the segmental interaction parameters. T = 300.15 K.

7.2.2.5 The phase behaviour of polymers in single solvents

The regressed segmental interaction parameters were now employed for the prediction of the LLE for the polymer-solvent systems, to detect phase splits around the temperature of interest. These intermediate results are given in Appendix G (§G.1). Table 7.4 summarizes the systems that gave phase splits, together with the composition of the phases in equilibrium at 300 K.

Table 7.4. Predicted compositions (%) of the phases in equilibrium for the systems polymer-solvent. w^{I}_{Pi} and w^{I}_{Pi} are the polymer weight fractions in phase I and phase II, respectively. The temperature is 300 K.

	***P*1**		***P*2**		***P*3**	
	w^{I}_{P1}	w^{II}_{P1}	w^{I}_{P2}	w^{II}_{P2}	w^{I}_{P3}	w^{II}_{P3}
A	2.02	99.85	0.0	38.56	0.0	99.99
B	0.0	81.53	miscible		0.0	88.71
C	miscible		miscible		0.0	80.78
D	miscible		0.0	65.83	0.0	48.68
E	miscible		0.0	95.03	0.0	64.19

7.2.3 Sub-task S1-V2.3: polymer in solvent binary mixture

As the FV-UNIQUAC does not require extra segmental interaction parameters for ternary systems, the phase behaviour for the ternary composition space was extrapolated from the binary interaction. Anyway, as no information on the ternary systems could be found, the predictions on the ternary composition space could not be validated.

When dealing with ternary systems, additional segmental interaction parameters for the FV-UNIQUAC model are required, to account for the solvent-solvent interactions. Solvent mixtures were previously analyzed (sub-task S1-V2.1, §7.2.1) employing non-polymer models such as UNIFAC. In the current task the interactions between solvents were modelled with the same model employed for the polymer-solvent systems (FV-UNIQUAC).

The information gathered and the results obtained in sub-task S1-V2.1 were then employed to fit the missing parameters. That is, if experimental data were available, they were used for the regression; otherwise the predicted solubility and/or phase equilibria were employed as pseudo-data for the regression.

Results are summarized in the following tables/figures:

- Table 7.5 summarizes information about the polymer-solvent systems under consideration. Information about the problems encountered during the regression for some of the systems under consideration ((*A*-*D*), (*A*-*E*)) are also shown. These problems are discussed in §7.2.3.1;
- Figs. 7.15-7.16 show the segmental interaction parameter regression results for the systems of Table 7.5 (systems (*A*-*D*) and (*A*-*E*) excluded);

- Table 7.6 lists the systems that show phase split and the compositions of the phases in equilibrium, according to the FV-UNIQUAC model.

Table 7.5. Type of experimental and pseudo-experimental data for the solvent-solvent systems, temperature dependence of the data, sum of square residuals, information about problems encountered during the parameter regression (last column).

nº	(1)	(2)	Experimental data	*T*-dependence	ΣR^2	Problems
1	*A*	*B*	pseudo x_i-ln(γ_i)	*T*-dep	6.34R-3	-
2	*A*	*C*	pseudo x_i-ln(γ_i)	*T*-dep	3.23E-3	-
3	*A*	*D*	pseudo x_i-ln(γ_i)	iso-*T*	4.06	yes
4	*A*	*E*	LLE data	-	-	yes
5	*B*	*C*	pseudo x_i-ln(γ_i)	*T*-dep	8.75E-3	
6	*B*	*D*	pseudo x_i-ln(γ_i)	*T*-dep	1.58E-4	-
7	*B*	*E*	pseudo x_i-ln(γ_i)	*T*-dep	7.15E-03	-
8	*C*	*D*	pseudo LLE data	*T*-dep	5.81E-08	-
9	*C*	*E*	pseudo x_i-ln(γ_i)	*T*-dep	1.42E-03	-
10	*D*	*E*	pseudo LLE data	*T*-dep	2.04E-05	-

Table 7.6. Predicted compositions (%) of the phases in equilibrium for the systems solvent-solvent. w^I_1 and w^{II}_1 are the solvent (1) weight fractions in phase I and phase II, respectively. The temperature is 300 K.

	A		***B***		***C***		***D***	
	w^I_A	w^{II}_A	w^I_B	w^{II}_B	w^I_C	w^{II}_C	w^I_D	w^{II}_D
A	-		-		-		-	
B	miscible		-		-		-	
C	miscible		miscible		-		-	
D	3.29	34.34	miscible		31.15	76.75	-	
E	*		miscible		miscible		5.11	30.26

*could not be calculated, see §7.2.3.1

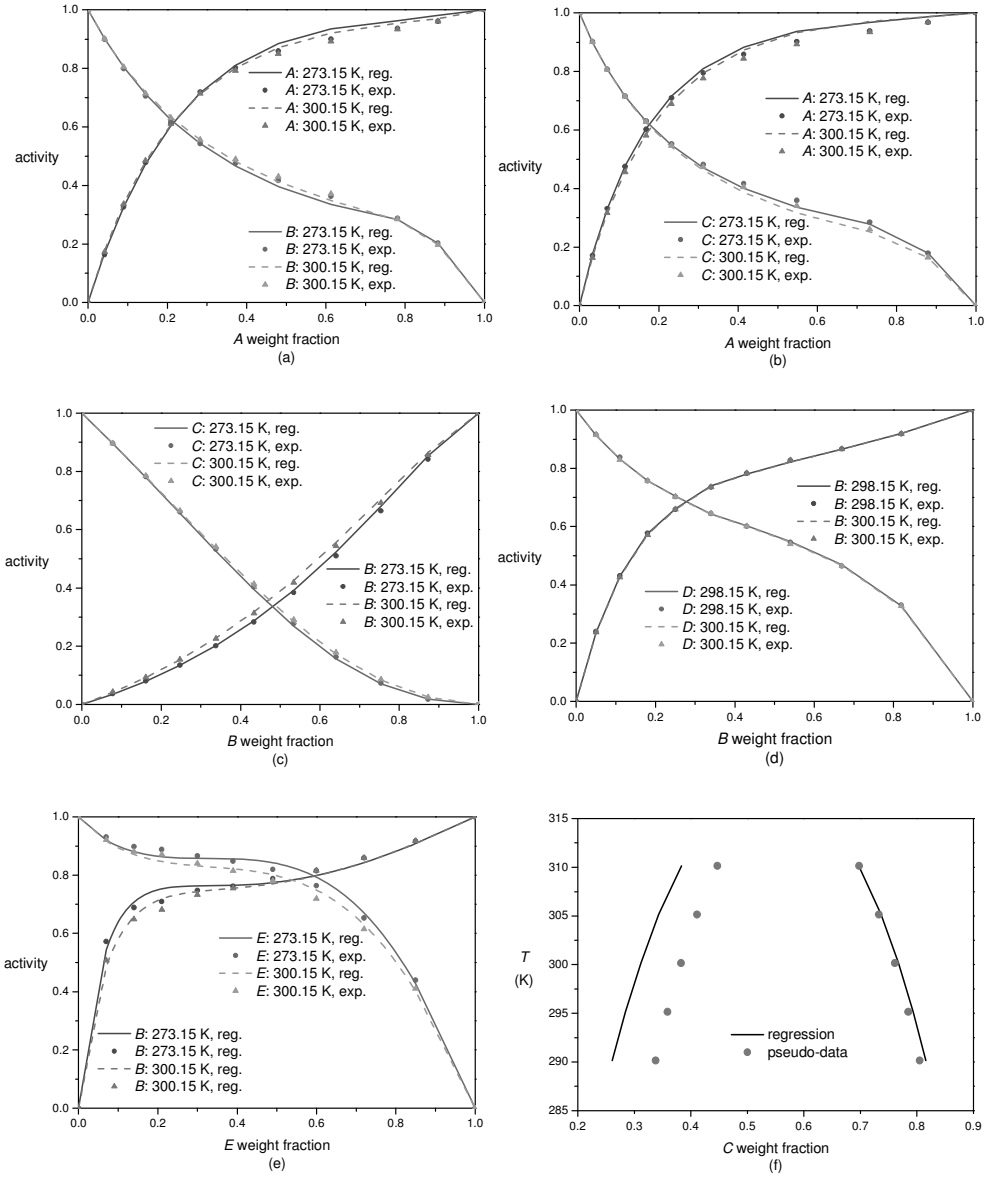


Figure 7.15. Regression of experimental (or pseudo-experimental) data for the systems: (a) (*A*-*B*); (b) (*A*-*C*); (c) (*B*-*C*); (d) (*B*-*D*); (e) (*B*-*E*); (f) (*C*-*D*).

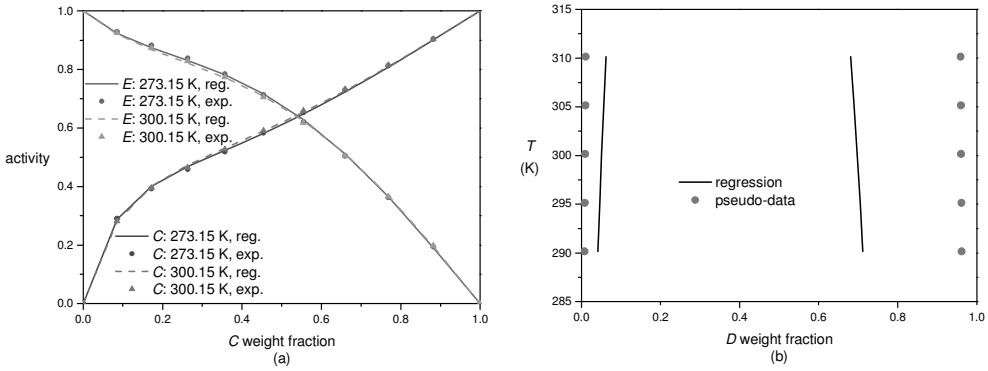


Figure 7.16. Regression of pseudo-experimental data for the systems (a) (*C*-*E*) and (b) (*D*-*E*).

7.2.3.1 Modelling problems solvent-solvent

The FV-UNIQUAC model was found not capable of describing the behaviour of the systems with strong intermolecular interactions, which are characterized by high number of contacts between the molecules of the two different chemical species. An explanation of the limitation of the model is given below.

In the UNIQUAC (Abrams and Prausnitz, 1975) model the (molar) activity coefficient γ_i is given by a combinatorial and a residual contribution. The combinatorial contribution corresponds to:

$$ln\,\gamma_i^C = ln\left(\frac{\varphi_i}{x_i}\right)+1-\frac{\varphi_i}{x_i}+\xi_i \tag{7.10}$$

ξ_i is the Staverman-Guggenheim correction (Sayegh and Vera, 1980) defined as:

$$\xi_i = \frac{z\cdot q_i}{2}\left[ln\left(\frac{\varphi_i}{\theta_i}\right)-1+\left(\frac{\varphi_i}{\theta_i}\right)\right] \tag{7.11}$$

For the meaning of the symbols, refer to Chapter 3 (§3.6-3.7).

In the FV-UNIQUAC the combinatorial and free-volume term are combined in the so called entropic-FV contribution as in Eq. 3.54 (Chapter 3), which is here recalled for clarity:

$$ln\,\gamma_i^{entr-FV} = ln\left(\frac{\varphi_i^{FV}}{x_i}\right)+1-\frac{\varphi_i^{FV}}{x_i} \tag{7.12}$$

The residual contribution to the activity coefficient (γ_i^R) is calculated in the same way in UNIQUAC and FV-UNIQUAC.

A comparison between Eqs. 7.10 and 7.12 reveals that the two models are structurally similar. They differ in the following points:

- The definition of the volume fraction, which is calculated from the molar volume in UNIQUAC and from the free volume (difference between the molar and the hard-core volume) in FV-UNIQUAC;
- The Staverman-Guggenheim contribution ξ_i, which is neglected in the FV-UNIQUAC.

For systems of molecules with large differences in size and free volume as in polymer-solvent systems, Eq. 7.12 is necessary in order to account for the free-volume effect. For systems of molecules with similar size and free volume, such as in the case of solvent-solvent systems, Eqs. 7.10 and 7.12 usually give similar results. In these cases the Staverman-Guggenheim contribution is also negligible. But there are some exceptions, such as the systems water-alkane. These systems are characterized by strong intermolecular bonds/high number of contact sites (the product $z \cdot q_i$ in Eq. 7.11 assumes high values) and by a value for the ratio φ_i/θ_i far from 1 (Larsen, 1986). It is clear that in these conditions the Staverman-Guggenheim contribution ξ_i is not negligible.

Fig. 7.17 shows the Staverman-Guggenheim contribution for the systems (*A*-*D*) and (*A*-*E*), and for other two systems for which the modelling with the FV-UNIQUAC model was successful ((*B*-*C*) and (*C*-*E*)). It can be noted that ξ_i assumes very high values for the systems (*A*-*D*) and (*A*-*E*), while for the systems (*B*-*C*) and (*C*-*E*) the contribution ξ_i is almost negligible. The ξ_i trend for the system (*A*-*E*) is expected, since it is a system water-alkane and this agrees with the observation of Larsen (1986).

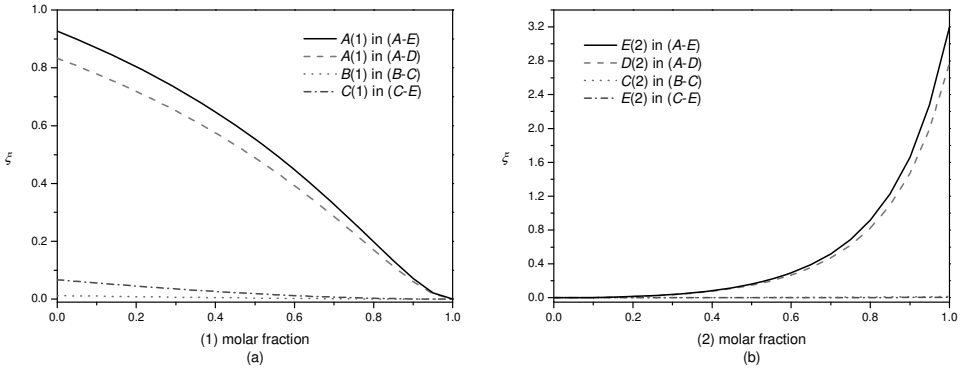


Figure 7.17. Staverman-Guggenheim contribution for compound 1 (a) and compound 2 (b) for the systems (*A*-*D*) and (*A*- *E*) and other two systems that did not show modelling problems with FV-UNIQUAC ((*B*-*C*) and (*C*-*E*)). T = 300 K.

The omission of the Staverman-Guggenheim contribution leads to the problem shown in Fig. 7.18. If LLE experimental data (pseudo-data for the system (*A*-*D*)) are regressed with the FV-UNIQUAC for the systems (*A*-*D*) and (*A*-*E*), and the interaction parameters are employed to predict the function delta Gibbs energy of mixing ($\Delta G^{mix}/RT$), two immiscibility regions are detected for both systems.

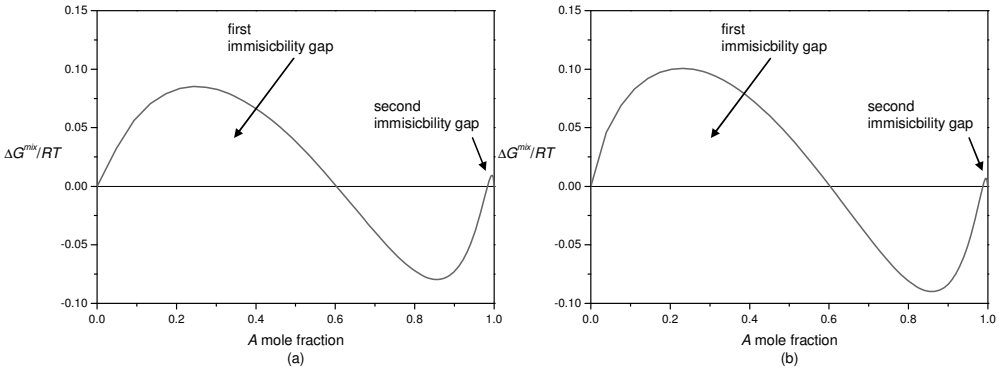


Figure 7.18. Function $\Delta G^{mix}/RT$ for the systems (a) (*A*-*D*) and (b) (*A*-*E*). The two immiscibility gaps detected for both systems are also highlighted. T = 300 K.

According to Fig. 7.18, two immiscibility regions exist (where the function $\Delta G^{mix}/RT$ takes positive values), while only one immiscibility region should be detected (according to the regressed data). If the activity coefficient trends are regressed instead of LLE (pseudo-) experimental data, the problem can be solved for the system (*A*-*D*), while the problem cannot be overcome for the system (*A*-*E*). Fig. 7.19 shows the trend of the function $\Delta G^{mix}/RT$ for the system (*A*-*D*), calculated with the interaction parameters obtained from the activity coefficient regression.

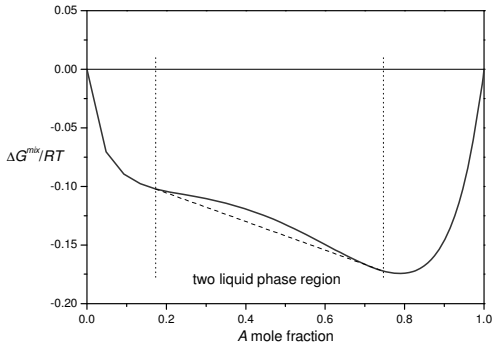


Figure 7.19. Trend of the function $\Delta G^{mix}/RT$ for the system (*A*-*E*), calculated with the interaction parameters regressed on the activity coefficients. T = 300 K.

In Fig. 7.19 the tangent plane is shown, and the following two phase region can be detected: x_A = 0.17-0.75 (for further explanations on the on the *TPD*, see Chapter 4, §4.1.3).
The fitting of the pseudo-experimental data for the system (*A*-*D*) was really poor (ΣR^2 is high, see Table 7.5), resulting in a poor prediction of the LLE.

7.2.3.2 The phase behaviour of polymer in solvent binary mixture

The regressed segmental interaction parameters were now employed for the prediction of the LLE for the ternary systems polymer-solvent binary mixture (T = 300 K). These intermediate results are collected in Appendix G (§G.2).

7.2.4 Sub-task S1-V2.4: polymer in multicomponent mixture

This task was not performed since the objective of the case study was to identify a pure solvent or to design a binary mixture of solvents that can dissolve the AI and match the *a priori* defined constraints. Only if no pure solvent or binary mixture would have satisfied the *a priori* defined criteria (§7.1.2) and/or did not dissolve the AI, ternary solvent mixtures would have been taken into consideration.

7.2.5 Sub-task S1-V2.5: copolymer in pure solvent & binary mixture

All the repeat units *M*1, *M*2 and *M*3 were here joined together to form the copolymer, so that the copolymer phase behaviour in single solvents and solvent binary mixtures could be simulated. The FV-UNIQUAC accounts also for the energetic interaction between the different monomers in a copolymer. Therefore, a new set of parameters was regressed: the segmental interaction parameters between the segments constituting the copolymer, (*M*1-*M*2), (*M*1-*M*3), (*M*2-*M*3). Fig. 7.20 shows the parameter regression results. Table 7.7 reports the necessary information for the monomer-monomer systems under consideration.

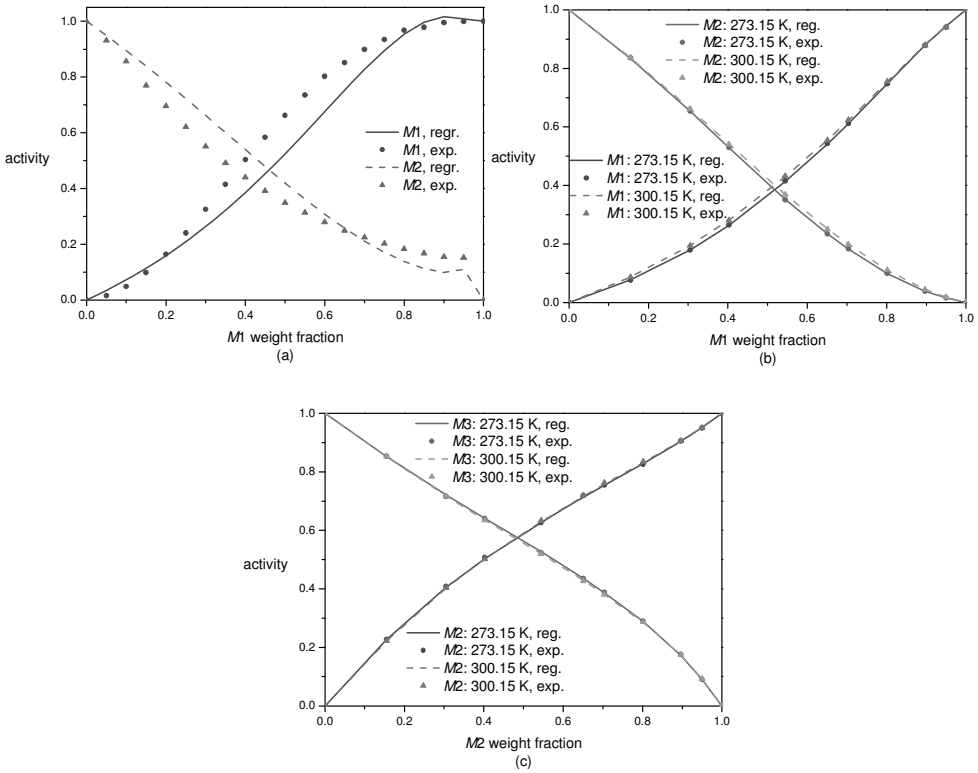


Figure 7.20. Regression of pseudo-experimental solubility data for the systems: (a) (*M*1-*M*2) at 300.15 K; (b) (*M*1-*M*3) and (c) (*M*2-*M*3) at two different temperatures.

Table 7.7. Type of experimental and pseudo-experimental data for the monomer-monomer systems, temperature dependence of the data, sum of square residuals, information about problems encountered during the parameter regression (last column).

nº	(1)	(2)	Experimental data	*T*-dependence	ΣR^2	Problems
1	*P*1	*P*2	pseudo x_i-ln(γ_i) (*M*1-*M*2)	iso-*T*	3.57	-
2	*P*1	*P*3	pseudo x_i-ln(γ_i) (*M*1-*M*3)	*T*-dep	1.72E-4	-
3	*P*2	*P*3	pseudo x_i-ln(γ_i) (*M*2-*M*3)	*T*-dep	1.21E-3	-

It can be noted that the regression of pseudo-experimental data for the system (*M*1-*M*2) was quite poor, also if the regression was performed using isothermal data.
At this point, all the necessary segmental interaction parameters were available and it was possible to proceed to task S1-V3, where the behaviour of the AI in solvents and binary solvent mixtures could finally be simulated and verified.

7.3 Task S1-V3: AIs verification

The goals of this task were:

- To simulate the phase behaviour of the copolymer in single solvents and binary solvent mixtures (sub-task S1-V3.1, §7.3.1);
- To analyze the effects of molecular weight (*Mw*) and molecular weight distribution (*Mwd*) on the copolymer phase behaviour. With this analysis it was possible to identify the conditions for which the two phase region has its maximum extension (for each system which shows phase split). In fact, to ensure the solubility of the polymer chains (that have different length since *PDI* ≠ 1) the worst scenario (the largest immiscibility gap) needed to be taken into consideration (sub-task S1-V3.2, §7.3.2);
- To apply the constraints of Eqs. 7.6-7.7, rejecting those systems that exhibit immiscibility for w_{Cop} = 0.05, independently from the solvent mixture composition (for instance, systems with all the three binary pairs showing large immiscibility gap) (sub-task S1-V3.3, §7.3.3).

7.3.1 Sub-task S1-V3.1: phase behaviour

The copolymer phase equilibria in single solvents were now calculated. The copolymer showed phase separation with solvent *A* and solvent *E*. Fig. 7.21 shows the relative phase equilibrium diagrams (*Cop* = copolymer). Table 7.8 shows the phase equilibrium composition for these two systems, at the temperature of interest (300 K).

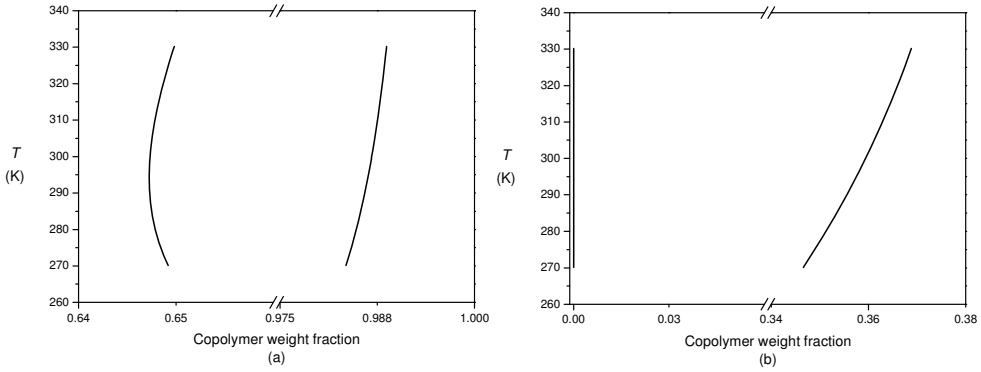


Figure 7.21. Predicted phase behaviour for the systems (a) (*Cop-A*) and (b) (*Cop-E*).

Table 7.8. Predicted compositions (%) of the phases in equilibrium for the systems copolymer-solvent. w_{Cop}, w_S are the copolymer and solvent weight fractions, respectively. T = 300 K.

System	Phase 1		Phase 2	
	w^I_{Cop}	w^{II}_S	w^I_{Cop}	w^{II}_S
Cop-A	64.7	35.27	98.68	1.32
Cop-E	0.00	100.0	35.94	64.06

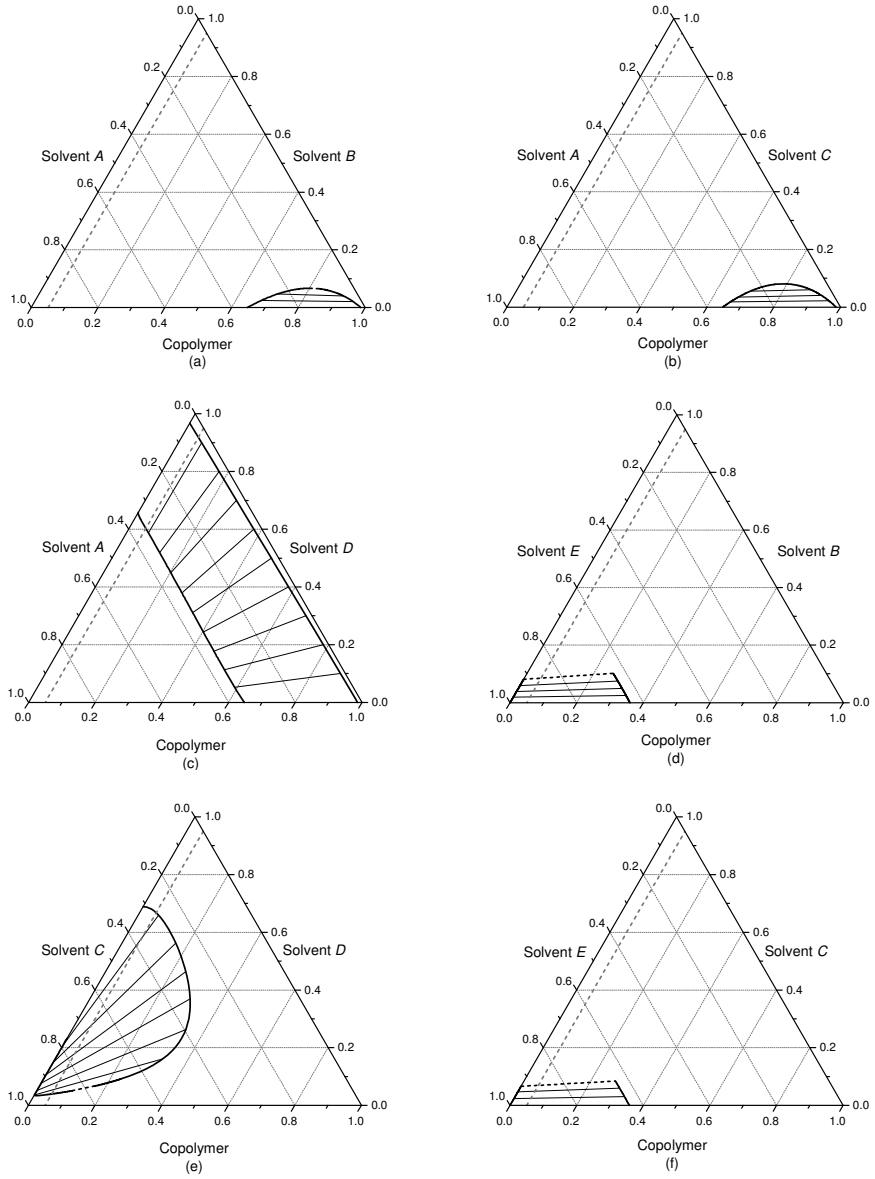


Figure 7.22. Weight based ternary phase equilibrium diagrams for the systems: (a) (*Cop-A- B*); (b) (*Cop-A-C*); (c) (*Cop-A-D*); (d) (*Cop-B-E*); (e) (*Cop-C-D*); (f) (*Cop-C-E*). The dashed line at w_{Cop} = 5% represents the copolymer composition in the formulation. T = 300 K.

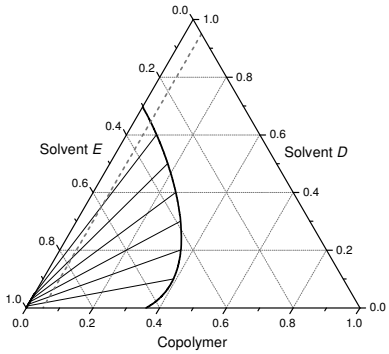


Figure 7.23. Weight based ternary phase equilibrium diagrams for the system (*Cop-D-E*). T = 300 K.

The phase behaviours of the copolymer in solvent binary mixtures were also calculated. In Figs. 7.22-7.23 the ternary phase diagrams are shown. The copolymer composition in the formulation (5% by weight) is also highlighted (red dashed line).
Since the interaction parameters for the system (*A-E*) could not be obtained, the ternary diagram for the system (*Cop-A-E*) could not be calculated. Note, however, that the system (*Cop-A-E*) would likely present a very large immiscibility region, since the copolymer shows immiscibility regions with both solvents *A* and *E*, and it was previously found that the system (*A-E*) shows large miscibility gap, too (see §7.2.1). Therefore the system (*Cop-A-E*) would unlikely have been a promising candidate formulation.

7.3.2 Sub-task S1-V3.2: property calculation

The objective here was to analyze the effects of molecular weight (*Mw*) and molecular weight distribution (*Mwd*) on the copolymer phase behaviour. The molecular weight distribution of the polymer was not known, but its variance is lumped in the parameter *PDI*, the polydispersity index. As the value of *PDI* increases, the *Mwd*-curve becomes wider, and the larger is the amount of polymer with short chain length.
With this analysis it was possible to identify the conditions for which the two phase region for the systems (*Cop-A*) and (*Cop-E*) have their maximum extension since the worst scenario (the largest immiscibility gap) should be taken into consideration to ensure the solubility of all the polymer chains. The given ranges for the *Mw* and *PDI* were: $55000 < Mw < 75000$; $3 < PDI < 3.5$.
The sensitivity analysis was performed on the binary systems (*Cop-A*) and (*Cop-E*). Figs. 7.24a and 7.24b show the effect of the molecular weight on the phase behaviour for a fix value of the *PDI* (base case value: $PDI = 3.5$). Figs. 7.24c and 7.24d show the effect of the *PDI* for a fix value of the *Mw* (base case value: $Mw = 65000$).

The ranges of values explored in the sensitivity analysis were larger than the given ones (55000 < *Mw* < 75000; 3 < *PDI* < 3.5) to be able to drawn general considerations on the polymer phase boundaries dependence on the *Mw* and *PDI*.

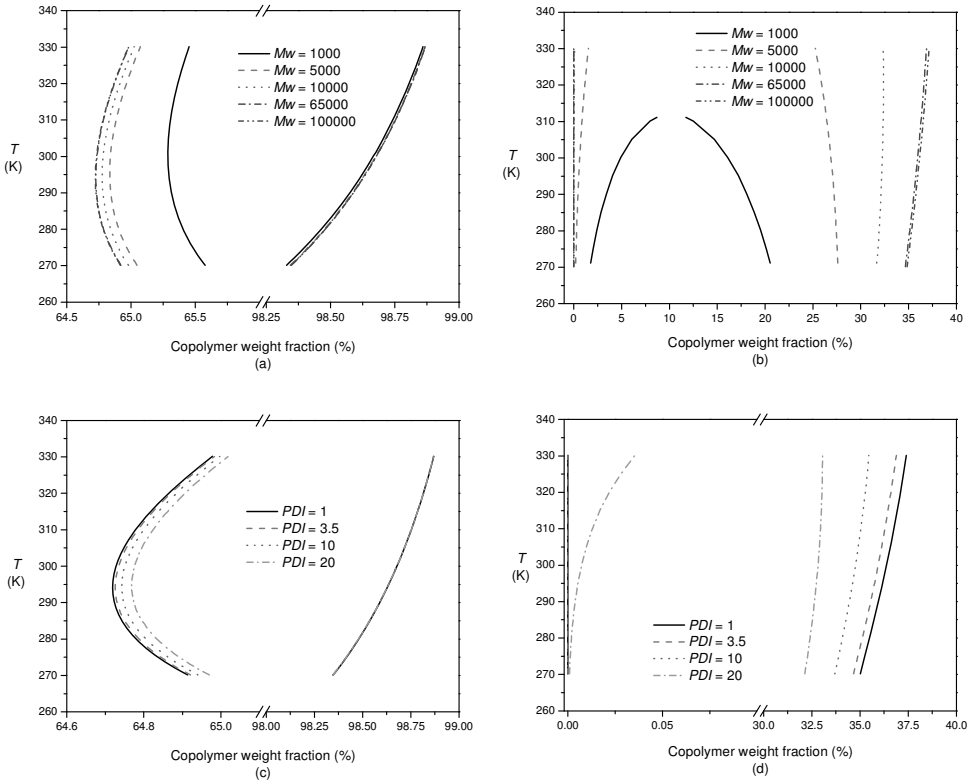


Figure 7.24. Effect of *Mw* on the copolymer phase behaviour for the systems: (a) (*Cop-A*); (b) (*Cop- E*). Effect of *PDI* on the copolymer phase behaviour for the systems: (c) (*Cop-A*); (d) (*Cop- E*).

In Figs. 7.24a and 7.24b the curves for *Mw* = 55000 and *Mw* = 75000 are not shown since they almost coincide with the trend for *Mw* = 65000, the base case. In Figs. 7.24c and 7.24d the trends for *PDI* = 3 and *PDI* = 3.5 are not also easily distinguishable. Therefore, it can be concluded that the systems under consideration are not really sensitive to small changes in *Mw* and *PDI*, such as the ranges under considerations. Therefore, in the next tasks of the verification the base case values for the molecular weight and the polidispersity index were taken into consideration (*Mw* = 65000; *PDI* = 3.5).

Interesting general observations can be drawn from Fig. 7.24:

- The two phase regions become smaller as the molecular weight decreases;

- The two phase regions become smaller as the *PDI* increases. In fact, increasing *PDI* means increasing number of polymeric chains with a lower molecular weight.

It can be concluded that, from the phase equilibria point of view, it is important to have a copolymer with low molecular weight and high polydispersity index. On the other hand, high molecular weight is necessary to guarantee the product functions (curl retention/holding power) and low *PDI* is required to guarantee the uniformity of the AI mechanical, physical and chemical properties that contribute to the product functions. Therefore, a compromise has to be found between the necessity of ensuring the product performance (AI effectiveness) and the phase stability. But this investigation of the AI functions goes beyond the scope of this case study (see the introduction to this chapter).

7.3.3 Sub-task S1-V3.3: screening

The objective of this sub-task was to apply the constraints of Eqs. 7.6-7.7 and, therefore, to reject those systems that show immiscibility for w_{Cop} = 5%. From Fig. 7.22 and 7.23 it can be noted and concluded that:

- The systems (*Cop*-*A*-*B*) and (*Cop*-*A*-*C*) are the only two systems showing one liquid phase at the copolymer composition of interest;
- All the other systems show miscibility issues at the copolymer composition of interest. But they are not rejected at this point, because the feasibility of the formula depends on the composition of the solvent mixture, which determines the position of the overall formula along the red dashed line.

7.4 Task S1-V4: solvent mixture verification

The binary solvent mixtures resulting from the combination of the solvents *A*, *B*, *C*, *D* and *E* are here checked against the target property constraints of Eqs. 7.1-7.5 defined in §7.1.2.

7.4.1 Sub-task S1-V4.1: property calculation

The mixture property models employed for the prediction of the mixture properties were linear mixing rules for the toxicity parameter and the dielectric constant, while rigorous models based on group contributions were selected to calculate the evaporation rate (Klein *et al.*, 1992) and the flash point (Liaw *et al.*, 2002) of the mixtures.
The MIXD algorithm was employed. It has to be underlined that the use of the MIXD algorithm for verification case studies is not absolutely necessary since the problem does not suffer of combinatorial explosion. In fact, the number of candidate solvent mixtures is only 9 (the total number of mixtures is 10, but the mixture (*A*-*E*) could be further considered).

Table 7.9 shows the results the mixtures matching the constraints, their composition in terms of weight fraction w_1, the cost (\$/kmol) and the values of the properties LC_{50} (mol/l), ε (-), T_{90} (s), T_f (K). The alcohol concentration (w_{OH}) matched the constraints for both the mixtures, too.

Table 7.9. Solvents mixtures that match the *a priori* defined targets: Eqs. 7.1-7.5.

nº	(1)	(2)	w_1	Cost	LC_{50}	ε	T_{90}	T_f
				\$/kmol	mol/l	-	s	K
1	*A*	*B*	0.645	26.11	0.706	70.0	945.95	303.5
2	*A*	*C*	0.597	20.29	0.637	70.0	849.85	302.1

7.4.2 Sub-task S1-V4.2: screening

In this sub-task the best performing candidate formulation needed to be selected. The candidate formulations with relative concentrations (weight based) are:

- (*Cop-A-B*): (0.05, 0.613, 0.337);
- (*Cop-A-C*): (0.05, 0.567, 0.383).

Fig. 7.25 shows the positions of the candidate formulations with respect to the phase boundaries. Both formulations fall in the single phase region. Therefore, both formulations are feasible candidates. The cost was chosen as performance index, and the cheaper mixture resulted to be (*A-C*). Consequently, the formulation (*Cop-A-C*) is cheaper.

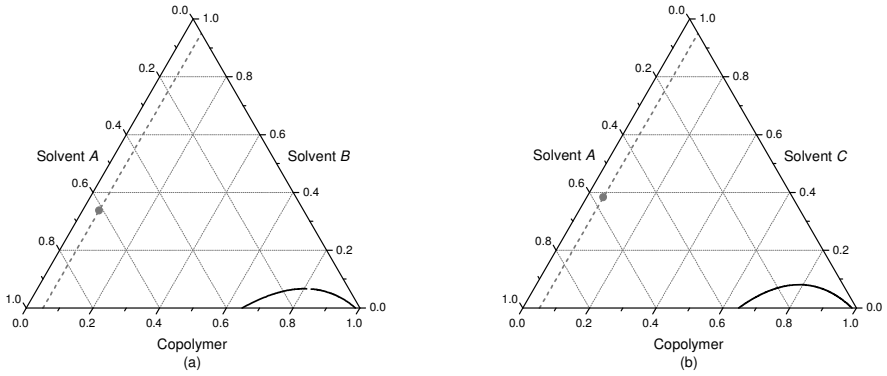


Figure 7.25. Position of the solvent mixture (●) with respect to the phase boundaries for the systems (a) (*Cop-A-B*); (b) (*Cop-A-C*). The T = 300 K.

7.5 Task S1-V5: additives verification

A neutralizing agent needs to be added in order to ensure that the product can be rinsed away with water. This task had two aims:

- To identify the minimum composition of the neutralizing agent that makes the copolymer itself water soluble. In fact, after the hair spray is applied on the hair, the delivery system (the solvent mixture (*A*-*C*)) evaporates, and the only component left on the hair is the copolymer. In order to remove the product from the hair with water (shower), the copolymer has to be water soluble (at least for high water concentrations).
- To ensure that the neutralizing agent (in the amount calculated above) is soluble in the formulation (*Cop*-*A*-*C*), with the composition calculated in §7.4.2 (w_{Cop} = 0.05; w_A = 0.567; w_C = 0.383).

This part was not performed for lack of time. But the two objectives above can be solved following the guidelines given in Chapter 5:

- Sub-task S1-V3.1: phase behaviour. The interactions of the neutralizing agent with the copolymer and water, together with the interactions between copolymer and water, have to be modelled with the FV-UNIQUAC model. Then, the phase behaviour of the ternary system constituted of the copolymer, water and the neutralizing agent has to be predicted. The minimum concentration of the neutralizing agent that makes the copolymer soluble in water needs to be identified. Then the interactions between the neutralizing agent and the solvents *A* and *C*, and the interactions between solvents *A* and *C* with water, have to be modelled with FV-UNIQUAC. The phase behaviour of this multicomponent system consisting of the copolymer, water, solvents *A* and *C*, and the neutralizing agent needs to be predicted.
- Sub-task S1-V5.2: property calculation. It is not required to verify the effect of the neutralizing agent on the overall formulation physical and chemical properties.
- Sub-task S1-V5.3: screening. If the system is single liquid phase at the desired composition of copolymer and neutralizing agent, the formulation is feasible and is the base case formulation on which the experiments should be performed. If the system is not a single liquid phase, the second best formulation (*Cop*-*A*-*B*) should be taken into consideration for the additive verification task.

7.6 Final remarks

Some attention should be focused on the reliability of the calculations/predictions performed in this case study. The main problem faced in the solution of this case study was the availability of experimental data: for very few systems polymer-solvent data could be found. That is why a purely predictive model was first considered (GC-Flory EoS), but it was found to be unable to predict the phase behaviour of oxygenated polymers in solvents. The FV-UNIQUAC model was therefore chosen, since it has been shown (Bogdanic and Vidal, 2000; Bogdanic, 2001) to handle

successfully the kind of systems under consideration in this case study. The problem of non availability of polymer-solvent experimental data has been overcome by applying a strategy that employs low molecular weight data (monomer-solvent) when polymer-solvent data are not available, or even using pseudo-experimental data for low molecular weight systems when experimental data can not be measured.
When using this strategy for the regression of the interaction parameters, it was obvious that the reliability of the predictions could have become questionable. In addition, it has been shown that the FV-UNIQUAC model cannot represent all the systems involved in the case study, and that for some of the systems the model performance is quite poor.
Keeping in mind all these drawbacks and that the alternative solution was to depend totally on having the time and resources to verify the system by experiments, it can be stated that almost everything that could be done for solving the problem, was tried and a reasonable and useful verification was possible. Also, the stated goals were satisfied.
The availability of experimental data for all the binary systems under consideration (15 polymer-solvent systems and 10 solvent-solvent systems), would of course, have strengthened the reliability of the results, leading to much more trustworthy results.

8

CONCLUSIONS

In this work a particular branch of product design has been investigated. Attention has been focused on a product class known as the 'consumer oriented chemicals based products' and, in particular, the formulated products with a liquid delivery system. Formulations are nowadays designed mainly through experiment-based trial-and-error techniques, with large investments of time and resources. This work proposed the use of computer-aided tools for a first screening stage, in which the number of alternatives is reduced to a reasonable sub-set of potential products, on which experimental techniques can be applied for testing and amendment.
In this chapter, the achievements of this work are summarized (§8.1), and some perspective on the challenges and future work is also given (§8.2).

8.1 Achievements

The issues and needs that were to be achieved in this work are listed as (see also Chapter 2, §2.6):

- Development of *property models*;
- Development of *methods* and *tools* for the solution of the CAM^{b}D problem and for the stability test of liquid mixtures, as well as databases and knowledge base;
- Development of a *methodology* for the design and verification of formulated products with a liquid delivery system;
- Development of systematic *framework*, based on the methodology, which includes the models, methods and tools, databases and knowledge base;
- Development of *case studies* to test and validate the methodology and the framework.

All the above issues and needs have been achieved in this work.

The modelling needs for the solution of the problem under consideration (design and verification of formulated products) have been established (Chapter 3). If accurate property models were available in the literature and all the needed interaction parameters were also available, the models were adapted in this work for the

estimations of the needed target properties. If the needed property models were not available and/or the accuracy was not satisfactory and/or the interaction parameters were not available, new property models were developed and/or new parameters were estimated.

M&G GC^+ *property models* were developed for the prediction of pure compound viscosity and surface tension. The M&G GC-based models showed very good statistics both in the regression ($SD < 1.5$) and in the prediction ($SD \sim 3.0$). The viscosity model was also compared with two other estimation models (Sastri and Rao, 1992; Orrick and Erbar, 1974): the M&G GC^+ model showed far better performance.

The CI-based models did not show as good test statistics as for the GC-based methods ($SD \sim 6.0$-7.0). The CI-based method alone does not provide improvement to the estimation of the surface tension or the viscosity, but it can be used as an auxiliary tool for the host GC-based method (the M&G method) to predict the missing group contributions. In fact, not more than 1-2 groups are usually missing for any molecular structure of an organic chemical, therefore the prediction of the property would be of acceptable accuracy with the combined GC-CI based method.

Linear correlations were also developed for the estimations of the evaporation time and cost of pure organic compounds.

Two new *methods* and the corresponding *tools* were developed in this work: a Mixture Design algorithm (MIXD) for the design of binary mixtures and a stability test algorithm (STABILITY), for the determination of the phase stability of liquid mixtures under consideration.

The newly developed algorithms (MIXD and STABILITY) were found to be essential tools in product design. The MIXD algorithm is a multi-level algorithm based on the reverse approach: given the desired property values, the algorithm identifies the binary solvent mixtures matching the constraints. The constraints are applied at three different levels: at first, linear constraints are applied, and the mixture composition minimizing the cost is calculated; then, non-linear constraints are employed to further screen the candidate mixtures; in the last level, the phase stability is checked through the phase stability algorithm, which is run directly from MIXD. The STABILITY algorithm is also a multi-level algorithm that employs three different levels of screening based on the analysis of the delta Gibbs energy of mixing, and identifies the total miscibility, partial miscibility (compositions of the two phases in equilibrium are calculated) and complete immiscibility.

Databases of AIs, solvents and additives were also developed, as well as knowledge base for supporting the decisions and choices performed during the design/verification of formulated products.

A systematic *methodology* that integrates computer-aided techniques and experimental validation was proposed. The methodology is composed by three main stages: a computer-aided stage where numerous alternatives are screened through the

use of modelling tools, a stage in which experiments are planned, and a last stage where the experiments are performed in order to test, verify and amend the base case product/products obtained at the end of the first stage.
Two different scenarios were considered: the design of completely new products and the verification of modified existing products. While the design involves a big combinatorial problem because the identities of the ingredients of the formulation are not known, the verification involves a small combinatorial problem since a shortlist of ingredients is suggested by the problem (user). In both cases, the goal is to identify the optimal formulation/formulations that also guarantee the phase stability. A fast screening methodology using short-cut models was developed for the design scenario, while a rigorous methodology was proposed for the verification scenario.

A systematic *framework* for the design and verification of formulated products was added to an existing software, the 'virtual Product-Process Design laboratory' (Morales-Rodriguez, 2009), extending the range of applicability of the software. The framework is based on the computer-aided stage of the systematic methodology for the design and verification of formulated products. It collects all the models, methods, tools, databases and knowledge base developed in this work, offering flexibility, user-friendliness, integration, data management, decomposition of complex problems, and so on.

The systematic methodology for the design and verification of formulated products, as well as the corresponding framework, were tested on a number of *case studies*. The case studies illustrated how the developed methodology is able to handle the complexity of products that are constituted of several ingredients (5 up to 20).
Three products (paint formulation, insect repellent lotion, sunscreen lotion) were considered for the design scenario, and two of these products (insect repellent lotion, sunscreen lotion) were also tested through the experimental stage. Results confirm that short-cut models can be employed for the screening of thousands of alternatives. Some modifications/improvements of the algorithm were proposed: at first, the Hansen solubility parameters should be added to the Hildebrand solubility parameter to ensure the solubility (mainly when dealing with solid solubility in liquids); secondly, more criteria should be taken into consideration during the computer-aided screening, such as the stability at different temperature than the design one (300 K), the flammability and so on.
A hair spray product was considered for the verification scenario. Here, it was shown how, following the proposed methodology, an industrial application can be decomposed, analyzed, and optimized. The main issue of the non availability of data was overcome with the generation of pseudo-experimental data with other models. The reliability of the results depends on the accuracy of the experimental data, and on the accuracy of the model employed for representing the system. The use of pseudo-experimental data compromised the trustworthiness of the results, but it was the only

alternative that was avilable, apart from performing numerous, time consuming and expensive experiments. A base case formulation was proposed, based on which the experimental planning and the experimental validation was started.

8.2 Challenges and future work

One of the challenges of product design lies in the problem definition (Harper, 2000; Gani, 2004a; Costa *et al.*, 2006). This work has tried to systematize this task, giving some ideas on how to identify the performance criteria, translate them into target properties and set the constraints. The translation of the performance criteria into physicochemical properties as well as the setting of the constraints are fields in which efforts should be focused. The main issue regarding the translation process concerns the choice of the properties associated to the particular performance criteria. Regarding the setting of the constraints, the main issue is related to the selection of the values range/target. In this work, information found in the literature have been employed, as well as information about patented products. The solution to this problem definition issue could be a kind of database in which, for every specific formulated product type (emulsions, tablets, sprays,…), performance criteria are collected, every performance criteria is related to a number of physicochemical properties, and value ranges/target values are suggested. This database could be shared on the net, and companies and laboratories would be able to share their knowledge through this platform.

Sensorial factors and cosmetic properties (appearance, turbidity, odour, skin feeling, stickiness,…) are very important for the type of products considered in this work, and they are of major importance for the success of the product on the market (as already underlined in Chapter 6, design case studies). It is still not clear how these factors are related to physical and chemical properties, therefore it is still not possible to consider them during the stage of computer-aided design of formulated products. Additional efforts should be focused in tracing the relation between these factors and the target properties, since their inclusion in the computer-aided design could spare time and resources during the experimental validation.

The relation between the product main functions (to repel the mosquitoes for an insect repellent, to block the UV radiations and/or reduce the risk of skin cancer for a sunscreen,…) and the concentration of the AIs in the formulation is another point on which future research could focus. In this work, information from literature and patents have been employed, but this has been proved to be ineffective in the case of a sunscreen (the target value for the SPF was 10-15, while a value of 6 was found through experimental validation).

One of the main limitations of the property models is the availability of the model (interaction) parameters. More effort should be focused in extending the parameter

tables in order to enlarge the application range of the models, and being able to apply them to a wider range of problems. A truly predictive modelling approach is necessary. This is the reason why the CI-based method supplies the GC-based method with model parameters when not all the group contributions are available for the compound under consideration.
The FV-UNIQUAC model is a powerful tool for systems that involves polymers, but a big effort is still required to obtain the model (interaction) parameters when they are not available. Before this work, parameters were available only for six polymers in few solvents. If the interaction parameters are available, the calculations are straightforward and the design/verification time can be significantly reduced.

This work proposed an alternative and more efficient way of solving formulation design and verification problems. An integrated approach was proposed, which considers, at first, the use of computer-sided tools to reduce the search space, and, later, an experimental stage to test and amend the product formula. One of its limitations is that it was developed and tested only on formulations with a liquid delivery system. Future work could consist of tackling the design and verification of other types of 'consumer oriented chemicals based products', such as emulsions (creams, dilute emulsions,…) or solid products (tablets, pastes, powders, granules,…) that are also very important and have wide applications. These products are even more challenging than formulations with a liquid delivery system, since they involve microstructures (dispersion of phases in emulsions, crystalline structure for powders and granules,…) which are determined not only by the physical and chemical properties, but also by the processing conditions (T, P,…).

The virtual lab extension was shown to be an efficient, flexible and user friendly tool for the solution of the formulation design and verification problems. The software now hosts four different tailor-made work-flows (including the formulations work-flow) for the solution of product and process design problems. But more work-flows could be added, enlarging the software applicability range and making it even more complete and useful. The number of virtual experiments that the user can perform should be higher, the number of compounds in the database should be larger, the number of models available should also be augmented, in order to be able to perform virtual simulations and experimentation of different scenarios and complexity conditions.

Appendices

A

Surface tension M&G GC$^+$: experimental data and predictions

The experimental data collected for the GC-CI surface tension model are given in Table A.1. Table A.2 gives the 8 data points for the compounds that cannot be represented by the M&G groups (these are the data points that are considered only in the regression for the atom contribution model).

Table A.1. 402 experimental data points employed for the parameter regression of the M&G surface tension model. Experimental data, GC and CI estimations.

nº	name	*exp*	GC *est*	CI *est*
1	Acetaldehyde	20.50	22.61	11.89
2	Acetanilide	43.90	43.90	41.10
3	Acetic acid	27.10	24.96	18.10
4	Acetic anhydride	31.93	31.93	27.02
5	Acetone	22.72	23.35	14.51
6	Acetonitrile	28.66	28.63	18.07
7	Acetophenone	39.04	35.80	27.89
8	Allyl acetate	25.80	25.01	23.49
9	Allyl alcohol	25.28	26.21	14.49
10	Aniline	42.12	38.95	32.52
11	Benzaldehyde	38.00	40.36	25.63
12	Benzamide	45.50	45.50	38.99
13	Benzene	28.22	29.03	18.49
14	Benzonitrile	38.79	38.13	31.44
15	Benzoylbromide	42.40	44.80	39.90
16	Benzoylchloride	38.60	39.44	33.21
17	Benzyl alcohol	34.80	34.80	27.25
18	Benzylamine	39.30	37.80	34.54
19	Benzyl benzoate	42.82	42.82	49.62
20	Biphenyl	39.20	39.20	34.29
21	Bromobenzene	35.24	34.89	35.85
22	1-Bromobutane	25.90	26.93	21.99
23	1-Bromo-4-chlorobenzene	37.50	38.45	45.04
24	1-Bromodecane	29.10	30.66	31.46
25	1-Bromododecane	30.40	31.90	34.61
26	Bromoethane	23.62	25.69	18.83
202	Heptanaldehyde	26.30	25.71	19.48
203	Heptane	19.66	19.17	14.62
204	Heptanoic acid	27.76	28.07	25.55
205	2-Heptanone	26.12	25.39	20.39
206	3-Heptanone	25.70	25.17	19.87
207	4-Heptanone	25.50	25.17	19.87
208	1-Heptene	19.80	20.09	14.76
209	Hexadecane	27.05	24.76	28.82
210	2,4-Hexadione	29.70	30.77	24.06
211	Hexane	17.89	18.55	13.04
212	Hexanedinitrile	45.45	45.45	34.84
213	Hexanenitrile	27.37	31.36	23.94
214	1-Hexanol	25.81	27.16	19.08
215	2-Hexanone	25.45	24.76	18.81
216	1-Hexene	17.90	19.47	13.19
217	Hexyl acetate	26.00	25.96	28.08
218	Iodobenzene	38.71	37.99	41.49
219	Iodoethane	28.46	27.58	22.49
220	1-Iodoheptane	30.00	30.69	30.38
221	1-Iodohexadecano	32.30	36.28	44.59
222	1-Iodohexane	29.50	30.06	28.81
223	Iodomethane	30.34	26.96	16.87
224	1-Iodo-3-methylbutane	28.10	28.33	28.43
225	1-Iodo-2-methylpropane	29.80	27.71	26.86
226	1-Iodooctane	30.20	31.31	31.96
227	1-Iodopentane	28.90	29.44	27.23

27	1-Bromohexane	27.40	28.18	25.15
28	Bromomethane	23.70	25.07	14.60
29	1-Bromo-3-methylbutane	25.60	26.44	24.86
30	1-Bromo-3-methylpropane	24.30	25.82	23.28
31	1-Bromonaphthalene	43.90	43.90	45.45
32	1-Bromononane	29.10	30.04	29.88
33	1-Bromopentane	26.90	27.56	23.56
34	p-Bromophenol	46.20	41.72	42.22
35	1-Bromopropane	25.26	26.31	20.41
36	2-Bromopropane	23.25	25.20	22.87
37	1-Bromotetradecane	30.80	33.15	37.78
38	o-Bromotoluene	34.20	34.94	39.51
39	p-Bromotoluene	33.90	34.39	39.51
40	Butanenitrile	26.92	30.11	20.78
41	1-Butanethiol	25.20	24.78	24.55
42	Butanoic acid	26.05	26.21	20.82
43	1-Butanol	24.93	25.92	15.93
44	2-Butanol	22.54	22.53	16.62
45	2-Butanone	23.97	23.52	15.65
46	2-Butoxyethanol	26.14	29.48	25.64
47	Butyl acetate	24.88	24.71	24.92
48	tert-Butyl acetate	21.90	21.93	27.48
49	Butylamine	23.44	25.67	23.14
50	sec-Butylamine	21.10	21.10	23.98
51	tert-Butylamine	16.87	17.21	25.99
52	Butylbenzene	28.70	29.17	25.94
53	sec-Butylbenzene	28.10	27.59	26.49
54	tert-Butylbenzene	27.70	27.39	28.27
55	Butyl butyrate	25.30	25.67	27.64
56	Butyl ethyl ether	20.20	20.32	19.59
57	Butyl formate	24.52	24.61	22.59
58	Butyl methyl ether	19.60	19.68	18.67
59	Butyl propionate	24.90	25.05	26.05
60	4-tert-Butylpyridine	33.10	33.41	38.35
61	Butyraldehyde	24.40	23.85	14.74
62	o-Chloroaniline	41.20	43.05	41.72
63	Chlorobenzene	32.99	32.61	27.69
64	1-Chlorobutane	23.18	24.19	16.55
65	2-Chlorobutane	21.60	21.15	17.91
66	1-Chlorododecane	29.30	29.16	29.17
67	1-Chlorohexane	25.73	25.43	19.70
68	Chloromethane	15.40	19.70	10.92
69	1-Chloro-3-methylbutane	22.80	23.69	19.41
70	1-Chloro-2-methylpropane	21.70	23.07	17.83
71	1-Chloronaphthalene	41.60	41.63	37.29
72	o-Chloronitrobenzene	45.20	45.12	51.44
73	m-Chloronitrobenzene	46.20	45.15	51.44
74	p-Chloronitrobenzene	43.20	44.57	51.44
75	1-Chloropentane	24.40	24.81	18.12
76	2-Chlorophenol	39.70	39.99	34.06
77	3-Chlorophenol	41.18	40.02	34.06
78	1-Chloropropane	21.30	23.57	14.97
228	1-Iodopropane	28.80	27.08	27.12
229	2-Iodopropane	26.60	27.08	27.12
230	p-Iodotoluene	36.80	37.49	45.15
231	Isobutyl acetate	23.06	23.60	26.13
232	Isobutylamine	21.75	24.55	24.42
233	Isobutylbenzene	27.00	28.05	27.15
234	Isobutyl butyrate	22.40	24.55	28.86
235	Isobutyl formate	23.30	23.49	23.87
236	Isobutyl propionate	26.10	23.93	27.27
237	Isobutyric acid	24.60	24.84	21.88
238	Isopentyl acetate	24.30	24.22	27.72
239	Isopentyl butyrate	25.10	25.17	30.43
240	Isopropyl acetate	21.76	22.04	24.26
241	Isopropylbenzene	27.69	26.96	25.20
242	Isopropyl formate	21.70	22.87	22.00
243	Methanethiol	23.90	23.90	18.50
244	Methanol	22.07	24.05	11.70
245	o-Methoxybenzaldehyde	42.60	41.69	34.60
246	p-Methoxybenzaldehyde	42.10	41.14	34.60
247	2-Methoxyethanol	30.84	27.62	21.56
248	o-Methoxyphenol	38.90	37.21	33.76
249	Methyl acetate	24.73	22.85	20.78
250	N-Methylaniline	36.90	35.60	34.48
251	Methyl benzoate	37.17	33.29	34.15
252	Methyl butanoate	24.62	23.81	23.49
253	2-Methyl-2-butanol	22.30	23.39	19.34
254	3-Methyl-1-butanol	23.71	25.42	18.72
255	2-Methyl-2-butene	17.15	17.29	19.34
256	Methyl acetoacetate	32.60	31.94	29.26
257	2-Methylbutyl acetate	24.30	24.34	27.42
258	3-Methylbutyric acid	25.10	25.71	23.68
259	Methylcyanoacetate	38.70	37.09	32.81
260	Methylcyclohexane	23.29	23.00	28.57
261	cis-2-Methylcyclohexanol	30.50	30.39	34.87
262	2-Methylcyclohexanone	31.50	31.70	32.72
263	3-Methylcyclohexanone	30.80	31.70	32.72
264	4-Methylcyclohexanone	30.50	31.70	32.72
265	Methylcyclopentane	21.72	18.77	24.38
266	Methyl decanoate	28.10	27.53	32.96
267	Methyl dichloroacetate	34.00	30.25	35.10
268	Methyl dodecanoate	29.20	28.78	36.12
269	Methyl formate	24.36	22.75	18.51
270	Methyl heptanoate	26.50	25.67	28.22
271	Methyl hexadecanoate	29.60	31.26	42.44
272	2-Methylhexane	18.80	18.06	15.83
273	3-Methylhexane	19.31	18.18	15.53
274	Methyl hexanoate	25.90	25.05	26.64
275	Methyl isobutyrate	23.20	22.85	24.55
276	1-Methylnaphthalene	37.60	37.57	31.76
277	2-Methyloctane	21.40	19.30	18.98
278	4-Methyloctane	21.90	19.42	18.69
279	Methyl octanoate	27.40	26.29	29.81

79	2-Chloropropane	19.16	20.40	16.62
80	3-Chloropropene	23.14	24.49	15.12
81	p-Chlorotoluene	32.20	32.11	31.35
82	o-Cresol	36.90	35.94	28.45
83	m-Cresol	35.69	35.96	28.45
84	p-Cresol	36.20	35.39	28.45
85	Cycloheptanol	32.70	36.97	35.47
86	Cyclohexane	24.16	25.35	24.99
87	Cyclohexanol	32.92	32.74	31.28
88	Cyclohexanone	34.57	34.05	29.14
89	Cyclohexene	26.17	28.64	22.80
90	Cyclohexylamine	31.22	31.22	39.01
91	Cyclopentane	21.88	21.12	20.79
92	Cyclopentanol	32.50	28.52	27.17
93	Cyclopentanone	32.80	29.83	25.02
94	Cyclopentene	22.20	24.42	18.68
95	p-Cymene	26.70	26.45	28.79
96	Decane	23.37	21.04	19.35
97	1-Decanol	28.51	29.64	25.39
98	1-Decene	23.60	21.95	19.49
99	Dibenzylamine	40.60	41.35	52.34
100	p-Dibromobenzene	39.30	40.72	53.20
101	1,2-Dibromoethane	39.55	35.32	31.01
102	Dibromomethane	39.05	34.70	29.44
103	1,2-Dibromopropane	33.90	34.95	34.76
104	Dibutylamine	24.12	23.67	29.70
105	Dibutyl ether	22.50	21.56	22.75
106	m-Dichlorobenzene	35.43	36.75	36.88
107	1,1-Dichloroethane	24.07	25.44	21.86
108	1,2-Dichloroethane	31.86	29.83	20.13
109	Dichloromethane	27.20	26.59	18.55
110	2,4-Dichlorophenol	43.50	42.60	43.25
111	1,2-Dichloropropane, (±)-	28.32	27.41	23.07
112	1,1-Diethoxyethane	20.89	21.10	26.55
113	Diethoxymethane	20.70	21.47	24.57
114	Diethylamine	19.85	21.19	23.38
115	N,N-Diethylaniline	34.00	35.38	39.28
116	Diethyl carbonate	25.90	26.97	29.02
117	Diethyl ether	16.65	19.08	16.43
118	Diethyl maleate	32.10	32.94	38.68
119	Diethyl malonate	31.30	31.65	37.43
120	Diethyl oxalate	31.60	31.75	35.56
121	Diethylsulfate	33.10	33.10	50.21
122	Diethyl sulfide	24.57	23.50	21.16
123	Diisobutylamine	21.72	21.44	32.20
124	Diisopentylamine	23.90	22.68	35.35
125	Diisopentyl ether	22.60	20.57	28.40
126	Diisopropylamine	19.14	18.90	28.59
127	Diisopropyl ether	17.27	17.03	21.50
128	1,2-Dimethoxybenzene	32.80	31.16	36.36
129	1,1-Dimethoxyethane	21.00	19.81	24.64
130	Dimethoxymethane	20.60	20.18	22.73
280	2-Methylpentane	16.88	17.43	14.24
281	3-Methylpentane	17.61	17.56	13.95
282	4-Methylpentanenitrile	26.60	30.24	25.15
283	Methyl pentanoate	25.30	24.43	25.07
284	2-Methyl-1-pentanol	25.00	26.16	20.00
285	2-Methyl-2-pentanol	22.90	22.35	21.27
286	3-Methyl-1-pentanol	25.00	26.16	20.00
287	3-Methyl-2-pentanol	24.90	24.13	20.63
288	3-Methyl-3-pentanol	23.30	22.35	20.83
289	4-Methyl-1-pentanol	24.10	26.04	20.30
290	Methylphenylsulfide	39.70	37.79	33.07
291	Methyl propanoate	24.44	23.18	21.91
292	2-Methyl-1-propanol	22.60	24.80	17.13
293	2-Methyl-2-propanol	19.96	21.45	18.56
294	1-Methylpropyl acetate	23.10	22.79	25.55
295	2-Methylpropyl acetate	23.10	23.60	26.13
296	2-Methylpyridine	33.00	33.00	32.24
297	3-Methylpyridine	34.50	34.50	32.24
298	4-Metylpyridine	34.90	34.59	32.24
299	N-Methyl-2-pyrrolidinone	40.21	40.21	38.84
300	Methyl salicylate	39.22	40.67	40.52
301	Methyl tetradecanoate	29.00	30.02	39.27
302	Naphthalene	40.10	40.10	28.09
303	Nitrobenzene	43.50	41.01	42.25
304	Nitroethane	32.13	30.80	30.30
305	Nitromethane	36.53	36.53	29.37
306	o-Nitrophenol	44.40	48.39	48.62
307	1-Nitropropane	30.10	31.43	31.88
308	2-Nitropropane	29.29	29.17	32.87
309	m-Nitrotoluene	40.80	41.09	45.91
310	o-Nitrotoluene	41.20	41.06	45.91
311	p-Nitrotoluene	39.80	40.51	45.91
312	Nonane	22.38	20.41	17.77
313	1-Nonanol	27.89	29.02	23.81
314	5-Nonanone	26.28	26.42	23.03
315	1-Nonene	22.60	21.33	17.92
316	Octadecane	27.87	26.01	31.97
317	Octane	21.14	19.79	16.19
318	1-Octane	21.30	19.79	16.19
319	Octanenitrile	27.60	32.60	27.10
320	1-Octanol	27.10	28.40	22.24
321	2-Octanol	25.90	25.02	22.94
322	Paraldehyde	25.63	25.97	42.75
323	Pentachloroethane	34.15	33.62	45.16
324	Pentadecane	26.70	24.14	27.24
325	Pentanal	25.44	24.47	16.33
326	2,4-Pentanedione	30.40	30.36	22.92
327	Pentanoic acid	26.70	26.83	22.40
328	1-Pentanol	25.36	26.54	17.51
329	2-Pentanol	23.45	23.16	18.20
330	2-Pentanone	23.25	24.14	17.22
331	3-Pentanone	24.74	23.93	16.71

131	Dimethylamine	26.34	26.34	21.11	332	cis-2-Pentene	16.80	17.37	11.71
132	N,N-Dimethylaniline	35.52	34.14	37.22	333	trans-2-Pentene	16.40	17.37	11.71
133	2,3-Dimethylbutane	16.90	17.67	15.38	334	Pentyl acetate	25.17	25.33	26.49
134	Dimethyl carbonate	28.60	25.72	27.10	335	Pentylamine	24.69	26.29	24.72
135	Dimethyl disulfide	33.39	32.85	25.17	336	Pentyl formate	25.50	25.23	24.16
136	2,4-Dimethylheptane	20.90	18.30	19.98	337	2-Phenylacetamide	44.30	44.30	41.23
137	2,5-Dimethylheptane	20.90	18.30	19.98	338	Phenylacetonitrile	41.70	41.70	33.69
138	2,6-Dimethylheptane	20.60	18.18	20.20	339	Phenylhydrazine	44.90	45.10	45.58
139	2,2-Dimethylpentane	17.60	17.32	17.61	340	Phenyl propyl ether	31.70	31.08	29.96
140	2,3-Dimethylpentane	19.50	18.41	16.67	341	Phenylsalicylate	42.80	42.80	31.38
141	3,3-Dimethylpentane	19.10	16.98	17.17	342	Phosphorus (III) chloride	27.98	27.98	21.37
142	2,5-Dimethylhexane	19.40	17.56	18.62	343	Phosphoryl chloride	32.03	32.03	24.47
143	2,4-Dimethylphenol	32.40	34.48	32.11	344	Piperidine	28.91	32.35	35.07
144	2,5-Dimethylphenol	34.60	34.48	32.11	345	Propanenitrile	26.75	29.49	19.20
145	3,5-Dimethylphenol	32.10	33.49	32.11	346	1-Propanethiol	24.20	24.16	22.98
146	Dimethyl sulfide	24.06	24.46	15.79	347	2-Propanethiol	21.33	21.21	24.85
147	Dimethyl sulfoxide	42.92	42.92	29.44	348	Propanoic acid	26.20	25.58	19.23
148	1,4-Dioxane	32.75	30.46	29.65	349	1-Propanol	23.32	25.29	14.34
149	Dipentyl ether	24.40	22.81	25.91	350	2-Propanol	20.93	21.79	15.33
150	Dipentyl sulfide	27.40	27.23	30.63	351	Propionamide	36.80	34.93	26.74
151	Diphenylamine	42.80	42.80	47.87	352	Propyl acetate	23.80	24.09	23.34
152	Diphenyl ether	26.75	26.75	40.33	353	Propylamine	21.75	25.05	21.56
153	Dipropoxymethane	22.80	22.71	27.72	354	Propylbenzene	28.50	28.55	24.36
154	Dipropylamine	22.31	22.43	26.54	355	Propyl benzoate	33.90	34.53	36.72
155	Dipropyl carbonate	26.40	28.21	32.17	356	Propyl butyrate	24.60	25.05	26.05
156	Dipropyl ether	20.00	20.32	19.59	357	Propyl formate	24.00	23.99	21.01
157	Dodecane	24.90	22.28	22.50	358	Propyl isobutyrate	23.30	24.10	27.05
158	1-Dodecanol	29.40	30.89	28.56	359	Propyl pentoate	25.30	25.67	27.64
159	Epichlorohydrin	36.36	36.93	24.21	360	Propyl propionate	24.20	24.43	24.48
160	1,2-Ethanediol	47.99	47.99	18.82	361	Pyridine	36.56	34.73	28.57
161	Ethanethiol	23.08	23.54	21.40	362	Pyrrole	37.06	34.72	26.58
162	Ethanol	21.97	24.67	12.76	363	Pyrrolidine	29.23	28.13	30.95
163	Ethanolamine	48.32	48.32	26.10	364	Quinoline	42.59	42.59	38.18
164	Ethoxybenzene	32.41	30.45	28.38	365	Succinonitrile	50.60	50.60	31.69
165	2-Ethoxyethanol	28.35	28.24	22.48	366	Sulfuryl chloride	28.78	28.78	45.15
166	Ethyl acetate	23.39	23.47	21.76	367	1,1,2,2-Tetrachloroethane	35.58	34.81	36.41
167	Ethyl acetoacetate	31.90	32.56	30.18	368	Tetrachloromethane	26.43	27.89	36.21
168	Ethylamine	19.20	24.43	19.98	369	Tetrachlorosilane	18.29	18.29	18.29
169	N-Ethylaniline	36.33	36.22	35.62	370	Tetradecane	26.13	23.52	25.67
170	Ethylbenzene	28.75	27.93	22.78	371	Tetradecanoic acid	31.60	32.42	36.60
171	o-Ethylbenzene	29.70	27.97	26.44	372	1-Tetradecanol	31.00	32.13	31.71
172	p-Ethylbenzene	28.30	27.42	26.44	373	1,2,3,4-Tetrahydronaphthalene	33.17	33.17	32.48
173	Ethyl butanoate	23.94	24.43	24.48	374	Thiophene	30.68	30.68	30.93
174	Ethyl chloroformate	26.20	27.55	27.08	375	Thymol	31.90	32.88	35.16
175	Ethyl crotonate	26.70	25.01	24.60	376	Toluene	27.73	28.56	22.15
176	Ethylcyanoacetate	36.10	37.71	33.81	377	p-Toluenesulfonylchloride	40.20	40.17	57.55
177	Ethylcyclohexane	25.15	24.58	29.20	378	m-Toluidine	37.90	39.02	36.18
178	Ethyl dodecanoate	27.90	29.40	37.10	379	p-Toluidine	37.20	38.44	36.18
179	Ethyl dichloroacetate	32.00	30.87	36.02	380	p-Tolunitrile	37.00	37.63	35.10
180	Ethyl formate	23.18	23.37	19.43	381	Tribromomethane	44.87	43.33	45.92
181	Ethyl hexadecanoate	30.70	31.88	43.42	382	1,2,3-Tribromopropane	44.80	44.57	46.65
182	Ethyl hexanoate	25.30	25.67	27.64	383	Tributylamine	24.39	24.20	36.65

183	Ethyl isobutyrate	22.70	23.47	25.47	384	1,1,1-Trichloroethane	25.18	24.25	30.90
184	Ethyl lactate	28.30	28.30	29.43	385	1,1,2-Trichloroethane	34.02	32.32	28.31
185	Ethyl 3-methyl butyrate	23.30	23.93	27.27	386	Trichloromethane	26.67	29.08	27.17
186	Ethyl methyl ether	15.30	18.44	15.52	387	Tridecane	25.55	22.90	24.09
187	Ethylmethylsulfide	24.40	25.08	18.47	388	1-Tridecene	25.80	23.82	24.24
188	3 Ethylpentane	20.00	18.18	15.24	389	Triethylamine	20.22	20.47	27.18
189	Ethyl pentanoate	24.70	25.05	26.05	390	Triethyl phosphate	29.50	29.50	43.67
190	Ethyl pentyl ether	21.70	20.94	21.16	391	Trifluoroacetic acid	13.53	13.29	22.82
191	Ethyl phenyl sulfide	36.50	38.41	35.75	392	Trimethylamine	13.41	13.41	24.14
192	Ethyl propyl ether	19.30	19.70	18.01	393	1,2,3-Trimethylbenzene	28.30	28.30	29.40
193	Ethyl propanoate	23.80	23.81	22.90	394	1,2,4-Trimethylbenzene	29.20	27.15	29.40
194	Ethyl thiocyanate	34.20	35.44	31.29	395	1,3,5-Trimethylbenzene	27.55	26.16	29.40
195	Fluorobenzene	26.66	28.17	20.27	396	2,2,3-Trimethylpentane	20.20	20.20	19.96
196	1-Fluorohexane	20.90	20.51	14.78	397	2,2,4-Trimethylpentane	18.40	16.82	20.40
197	1-Fluoropentane	19.50	19.89	13.21	398	Tripropylamine	22.40	22.34	31.92
198	m-Fluorotoluene	29.20	28.24	23.93	399	o-Xylene	29.76	28.60	25.81
199	p-Fluorotoluene	27.70	27.67	23.93	400	m-Xylene	28.47	28.63	25.81
200	Formanilide	42.10	42.10	38.77	401	p-Xylene	28.01	28.05	25.81
201	Furfural	43.09	43.09	26.03	402	Undecane	24.21	21.66	20.92

Table A.2. Two experimental data points that cannot be represented by the M&G groups. They were added to the list of compounds of Table A1 for the atom parameter regression of the M&G surface tension model. Experimental data and CI estimations.

n°	**name**	***exp***	**CI-*est***
1	Phenyl isothiocyanate	40.00	48.93
2	Tribenzylamine	40.00	70.70

B

Viscosity M&G GC$^+$: experimental data and predictions

Except the book of Viswanath and Natarajan (1989) and the book of Weast (1984), the experimental data for the development of the GC-CI viscosity model were collected from the following sources: Knapstad *et al.*, 1989; Mundhwa *et al.*, 2006; Cruz *et al.*, 2002; Martin *et al.*, 2001; Lafuente *et al.*, 1996; Ripple and Defibaugh, 1997; Roy *et al.*, 2007; Rodríguez, 1997; Das and Roy, 2006; Gascón *et al.*, 2000; Ćwiklińska *et al.*, 2007; Ali *et al.*, 2007; Yang *et al.*, 2007; Al-Hayan and Abdul-latif, 2006; Wang *et al.*, 2007; Kinart *et al.*, 2006; Li *et al.*, 2007; Al-Kandary *et al.*, 2006; Giner *et al.*, 2006; Wei and Rowley, 1984.

The 430 experimental data used for the development of the viscosity model are given in Table B.1. Table B.2 gives the 8 data points for the compounds that cannot be represented by the M&G groups (these are the data points that are considered only in the regression for the atom contribution model).

Table B.1. 430 experimental data points employed for the parameter regression of the M&G viscosity model. Experimental data, GC and CI estimations.

n°	name	*exp*	GC *est*	CI *est*	n°	name	*exp*	GC *est*	CI *est*
1	Acetaldeyde	0.21	0.33	0.96	216	3-Fluorophenetole	1.15	1.10	2.00
2	Acetic acid	1.03	1.12	1.21	217	4-Fluorophenetole	1.18	1.10	2.39
3	Acetic anhydride	0.82	0.82	1.32	218	2-Fluorotoluene	0.62	0.64	1.24
4	Acetone/dimethyl ketone	0.30	0.30	0.79	219	3-Fluorotoluene	0.56	0.58	1.24
5	Acetonitrile	0.35	0.29	0.93	220	4-Fluorotoluene	0.57	0.59	1.24
6	Acetyl chloride	0.36	0.46	0.76	221	Formamide	3.18	1.00	1.45
7	Allyl alcohol/2-propen-1-ol	1.17	1.65	1.19	222	Formic acid	1.55	3.42	1.43
8	Allyl bromide3-bromopropylene	0.46	0.40	1.33	223	n-Formylmorpholine	7.28	3.33	4.37
9	Allylchloride/3-chloropropene	0.31	0.30	0.85	224	2-Furaldehyde/furfural	1.44	1.44	2.72
10	Allyl ether/diallyl ether	0.41	0.30	1.26	225	Glycerin trinitrate	25.73	20.10	18.96
11	Allyl iodide/3-iodopropene	0.68	0.59	0.78	226	Heptadecane	3.40	3.10	3.43
12	Allyl thiocyanate	0.67	0.68	1.34	227	Heptane	0.38	0.37	1.17
13	2-((2-aminoethyl)amino)ethanol	88.20	88.20	2.66	228	2-Heptanethiol	0.82	0.83	0.77
14	n-Amyl acetate/n-pentyl acetate	1.22	0.81	1.57	229	Heptanoic acid/heptylic acid	3.68	3.25	2.19
15	Tert-amyl alcoholdimethylethylcarbinol,2-methyl-2-butanol	3.42	3.50	0.81	230	1-Heptanol/n-heptyl alcohol	5.53	4.73	1.81

16	Amyl amine/pentyl amine	0.68	0.90	1.50	231	2-Heptanol	3.70	3.34	1.45
17	Amyl butyrate/pentyl butyrate	1.09	1.04	2.05	232	2-Heptanone/methyl n-hexyl ketone	0.68	0.73	1.27
18	2-Amyl undecanol	27.78	27.96	3.82	233	1-Heptene	0.33	0.30	1.19
19	Aniline	3.64	2.75	2.58	234	Heptyl acetate	1.09	1.24	1.94
20	Anisole	1.03	0.92	2.08	235	Heptylamine	1.27	1.38	1.87
21	Benzene	0.59	0.58	2.28	236	Heptyl mercaptan	0.96	0.93	1.09
22	Benzonitrile/phenyl cyanide	1.23	1.38	2.69	237	2-Heptyl nonanol	28.05	27.96	3.82
23	Benzophenone	12.55	11.55	6.50	238	Hexadecane	2.92	2.51	3.08
24	Benzyl alcohol	5.17	5.17	2.66	239	2-Hexadecanol	29.17	22.62	3.79
25	Benzylamine	1.57	1.57	2.72	240	Hexadecyl acetate/cetyl acetate	6.38	8.36	5.14
26	Benzyl benzoate	7.88	7.88	8.25	241	2-Hexadecyl acetate	5.77	5.76	4.11
27	Benzyl chloride/alpha-chlorotoluene	1.24	1.24	1.90	242	Hexane	0.30	0.30	1.05
28	Benzylcyanide	1.89	1.89	2.76	243	Hexanenitrile/capronitrile	0.89	0.86	1.51
29	Benzyl ether	4.42	4.42	6.29	244	1-Hexanethiol/hexyl mercaptan	0.75	0.75	0.97
30	Bisdimethylaminophosphoryl chloride	3.42	2.56	1.06	245	2-Hexanethiol	0.64	0.67	0.70
31	Bromoethane/ethyl bromide	0.37	0.40	1.19	246	1-Hexanol	4.26	3.82	1.63
32	1-Bromo-2-methylpropane	0.59	0.54	1.09	247	N-Hexylamine	1.88	1.12	1.68
33	3-Bromoaniline/m-bromoaniline	5.40	5.07	1.87	248	Hexyl benzene	1.56	1.61	2.95
34	Bromobenzene	1.05	1.08	1.65	249	Hexyl ethanoate/hexyl acetate	1.03	1.00	1.75
35	1-Bromodecane	3.23	2.19	2.80	250	1-hexyne	0.35	0.35	1.05
36	1-Bromo-2,2-difluoroethane	0.72	0.59	1.08	251	Hydrogen cyanide	0.18	0.29	0.93
37	bromoform	1.81	2.75	1.64	252	Indan	1.32	1.32	3.75
38	1-Bromopropane/propyl bromide	0.48	0.49	1.32	253	Indene	1.61	1.61	3.52
39	2-Bromopropane/isopropyl bromide	0.45	0.44	0.86	254	Iodobenzene	1.51	1.51	0.76
40	2-Bromopyridine	1.70	1.53	2.00	255	Iodoethaneethyl iodide	0.55	0.59	0.69
41	2-Bromotoluene/o-bromotoluene	1.35	1.22	1.12	256	Iodomethane/methyl iodide	0.46	0.47	1.00
42	3-Bromotoluene/m-bromotoluene	1.13	1.11	1.12	257	1-Iodo-2-methylpropane	0.81	0.80	0.64
43	4-Bbromotoluene/p-bromotoluene	0.79	1.12	1.12	258	1-Iodopropane/propyl iodide	0.69	0.73	0.76
44	Butane	0.16	0.20	0.85	259	2-Iodopropane/isopropyl iodide	0.64	0.64	0.46
45	1-Butanethiol/butyl mercaptane	0.50	0.49	0.78	260	Isoamylamine/isopentyl amine	0.65	0.80	1.12
46	2-Butanethiol/1-methyl-1-propanethiol	0.47	0.44	0.56	261	Isoamylcyanide/isocapronitrile	0.88	0.76	1.13
47	1-Butanol/butyl alcohol	2.42	2.50	1.31	262	Isoamyl ether/isopentyl ether	0.91	0.82	1.05
48	2-Butanol/sec-butyl alcohol	2.88	1.77	1.05	263	Isobutyl acetate	0.65	0.58	1.06
49	1,3-Butadiene	0.14	0.13	0.89	264	Isobutylamine	0.56	0.65	1.01
50	Butyl acetate	0.67	0.65	1.40	265	1-(Isocyanato-1 methylethyl)-3-(1-methylethenyl)benzene	3.00	3.10	1.95
51	Butylamine/1-butanamine	0.56	0.73	1.35	266	Isoheptane/dimethyl butyl methane	0.35	0.33	0.88
52	Butylbenzene	0.93	1.05	2.37	267	3-Isopropyl-2-oxazolidinone	3.49	3.49	2.64
53	Butyl butyrate	0.91	0.84	1.84	268	Isovaleraldehyde/3-methylbutanol	0.51	0.56	1.04
54	2-Butyldodecanol	28.42	27.96	3.82	269	2-Methoxyethanol	1.49	1.34	1.26
55	Butyl ether/dibutyl ether	0.63	0.68	1.52	270	2-(2-Methoxyethoxy)ethanol	3.42	2.88	1.83
56	Butyl formate	0.63	0.60	1.66	271	2-[2-(2-Methoxyethoxy)ethoxy]ethanol	5.94	6.58	2.65
57	7-Butyl-1-hexyldecahydronaphthalene	17.46	15.97	7.44	272	N-2-Methoxyethyl acetamide	10.09	10.09	1.47
58	2-Butyl-3-hexylnaphthalene	17.98	17.98	5.37	273	1-Methoxypropane/methyl-n-propyl ether	0.24	0.27	0.91
59	3-Butyl-2-oxazolidone	4.68	4.40	3.74	274	Methyl acetate	0.36	0.35	0.95
60	3-Tert-butyl-2-oxazolidinone	5.34	5.95	2.03	275	Methylamine	0.18	0.39	0.94
61	Butyl phenyl ether	1.65	1.73	3.12	276	2-Methylaminoethanol	9.77	9.77	1.43
62	Butyl propionate	0.76	0.68	1.66	277	N-Methylaniline	1.96	1.55	2.28
63	Butyl valerate	1.10	1.04	2.05	278	2-Methylbenzenamine/o-toluidine	3.61	3.12	1.75
64	Bbutyraldehyde	0.42	0.51	1.24	279	3-Methylbenzenamine/m-toluidine	3.23	2.83	1.75
65	Butyric acid	1.38	1.72	1.58	280	Methylbenzene/toluene	0.54	0.60	1.55

66	Butyric anhydride	1.40	1.48	2.25	281	Methylbenzoate	1.79	1.62	2.73
67	Butyronitrile	0.54	0.56	1.22	282	2-Methylbenzonitrile/o-tolunitrile	1.54	1.57	1.83
68	Butyryl chloride	0.53	0.65	0.99	283	3-Methylbenzonitrile/m-tolunitrile	1.45	1.42	1.83
69	Carbon tetrachloride	0.88	0.86	0.39	284	4-Methylbenzonitrile/p-tolunitrile	1.51	1.43	1.83
70	Chloral	1.23	1.23	0.64	285	2-Methyl-1,3-butadiene/isoprene	0.20	0.19	0.73
71	2-Chloroaniline	3.17	3.98	1.64	286	2-Methylbutane/isopentane	0.21	0.24	0.94
72	1-Chlorodifluoroethane	0.30	0.25	0.46	287	2-Methyl-2-butanethiol/	0.60	0.59	0.41
73	Chlorodifluoromethane	0.18	0.22	0.59	288	2-Methyl-1-butanol	4.16	2.70	1.13
74	Chlorofluoromethane	0.24	0.40	0.70	289	2-Methyl-1-butanol (optiacally inactive)/isoamyl alcohol	3.49	2.74	1.09
75	Chloroform	0.53	0.71	0.57	290	3-Methyl-2-butanol	3.61	2.91	0.88
76	2-Chloro-6-methylaniline	2.14	3.19	1.12	291	2-Methyl-2-butene/beta-isoamylene	0.20	0.22	0.57
77	2-Chloro-6-methylpropane	0.42	0.40	0.70	292	Methyl butyrate/methyl butanoate	0.53	0.44	1.24
78	3-Chlorophenol/m-chlorophenol	10.68	10.33	1.68	293	Methyl chloride/chloromethane	0.26	0.20	0.76
79	2-Chlorophenylmethyl ether	2.00	2.12	2.49	294	2-Methyl-2-chloropropane	0.46	0.50	0.44
80	1-Chloropropane/propyl chloride	0.33	0.37	0.84	295	Methylcyclohexane	0.66	0.68	1.88
81	2-Chloropropane	0.30	0.30	0.60	296	3-Methylcyclohexanol	16.42	16.42	2.19
82	2-Chloropyridine	1.15	1.09	1.76	297	3-Methylcyclohexanone	1.54	1.29	2.10
83	2-Chlorotoluene/o-chlorotoluene	0.94	0.87	0.98	298	Methylcyclopentane	0.47	0.72	1.59
84	3-Chlorotoluene/m-chlorotoluene	0.80	0.79	0.98	299	Methyldiphenyl amine	5.67	9.54	4.43
85	4-Chlorotoluene/p-chlorotoluene	0.82	0.80	0.98	300	Methylene bromide	0.96	1.01	1.82
86	m-Cresol	11.74	8.09	1.80	301	2-Methyl-6-ethyl aniline	3.18	2.85	1.48
87	o-Cresol	7.28	8.92	1.80	302	Methyl ethyl ketone/2-butanone	0.39	0.38	0.92
88	p-Cresol	13.21	8.15	1.80	303	Methyl formate	0.32	0.32	1.12
89	Cumene/isopropylbenzene	0.72	0.72	1.68	304	2-Methyl-5-heptanol/6-methyl-3-heptanol	1.35	3.66	1.24
90	2-Cyanopyridine	1.92	1.96	3.27	305	3-Methyl-2-heptanol/5-methyl-6-heptanol	4.52	5.43	1.26
91	1,3-Cyclohexadiene	0.55	0.58	2.40	306	3-Methyl-6-heptanol/5-methyl-2-heptanol	2.10	3.61	1.24
92	Cyclohexane	0.87	0.71	2.74	307	3-Methylhexane	0.34	0.32	0.91
93	Cyclohexanone	1.95	1.34	3.07	308	2-Methyl-2-hexanol	3.82	5.35	1.00
94	Cyclohexene	0.61	0.64	2.57	309	Methyl isobutyrate	0.48	0.34	0.95
95	Cyclohexylamine	1.86	1.86	3.11	310	2-Methyl octane-2-thiol	1.14	1.38	0.63
96	Cyclohexyl bromide	2.10	1.74	2.04	311	Methyl oleate	4.79	5.87	5.00
97	Ccyclohexyl chloride	1.42	1.42	1.77	312	3-Methyl-2-oxazolidinone	2.41	2.30	2.41
98	Cyclopentane	0.40	0.75	2.33	313	3-Methyl pentane	0.30	0.26	0.79
99	Ccis-decahydronaphtalene	2.92	3.47	5.40	314	4-Methyl-2-penatanone	0.53	0.52	0.86
100	Decane	0.82	0.70	1.61	315	2-Methyl-1-pentene	0.27	0.29	0.78
101	1-Decanol/decyl alcohol	10.16	8.94	2.50	316	2-Methyl-2-propanethiol	0.57	0.52	0.35
102	Decene	0.75	0.57	1.65	317	2-Methyl-1-propanol/iso-butyl alcohol	3.15	2.21	0.98
103	1,2 Diaminopropane	1.47	1.25	1.50	318	2-Methyl-2-propanol/tert-butyl alcohol	3.90	3.06	0.69
104	1,2-Dibromoethane	1.55	1.25	2.05	319	Methyl propionate	0.42	0.36	1.12
105	1,2-Dibromoethylene	0.89	0.89	1.37	320	2-Methylpropionic acid	1.19	1.53	1.21
106	1,2-Dibromo-2-methylpropane	1.93	2.15	1.06	321	Methyl alculat/dimethyl sulfide	0.28	0.30	1.06
107	1,2 Dibromopropane	1.50	1.35	1.53	322	3-Methyl sulfolane	11.11	9.76	2.17
108	Dibutylamine/n-butyk carbonate	0.89	1.06	1.68	323	Methyl sulfoxide/dimethyl sulfoxide	1.92	2.22	0.90
109	Dibutyl carbonate	1.68	1.81	2.50	324	Methyl thiocyanate	0.73	0.67	1.21
110	Dichloroacetic acid	4.70	3.98	1.15	325	Nitrobenzene	1.81	1.96	4.31
111	1,2-Dichlorobenzene	1.26	1.11	0.93	326	2-Nitro-2-butene	0.84	0.79	1.41
112	1,3-Dichlorobenzene	1.00	1.01	0.93	327	2-Nitro-2-heptene	1.40	1.42	2.02
113	1,4-Dichlorobutane	1.27	1.05	1.02	328	3-Nitro-3-heptene	1.35	1.34	2.13
114	1,1-Dichloro-2,2-difluoroethane	0.65	0.57	0.54	329	3-Nitro-2-hexene	1.14	1.15	1.83
115	Dichlorodifluoromethane/ alcu-12	0.25	0.34	0.45	330	3-Nitro-3-hexene	1.04	1.08	1.90
116	1,1-Dichloroethane	0.46	0.45	0.58	331	3-Nitro-2-nonene	2.16	2.17	2.54
117	1,2-Dichloroethane	0.76	0.68	0.83	332	5-Nitro-4-nonene	2.04	2.05	2.63

118	Trans-1,2-dichloroethylene	0.38	0.38	0.67	333	3-Nitro-2-octene	1.76	1.75	2.28
119	Dichlorofluoromethane	0.31	0.26	0.58	334	4-Nitro-4-octene	1.69	1.66	2.36
120	Dichloromethane	0.41	0.46	0.75	335	2-Nitro-2-pentene	0.90	0.93	1.62
121	1,2-Dichlorotetrafluoroethane	0.36	0.48	0.37	336	2-Nitrotoluene/o-nitrotoluene	2.09	2.23	2.93
122	Diethylamine	0.31	0.45	1.09	337	3-Nitrotoluene/m-nitrotoluene	2.03	2.02	2.93
123	N,N-Diethylaniline	1.91	1.92	2.51	338	4-Nitrotoluene	1.26	2.03	2.93
124	2,6-Diethylaniline	3.52	3.26	1.84	339	Nonadecane	4.77	4.74	4.27
125	Diethyl carbonate/ethyl carbonate	0.74	0.77	1.63	340	Nonane	0.65	0.57	1.45
126	Diethylen glycol dinitrate	6.36	8.08	9.52	341	1-Nonanethiol/nonyl mercaptan	1.49	1.42	1.34
127	Difluoroacetic acid	2.52	1.68	1.19	342	2-Nonanethiol	1.32	1.27	0.97
128	2,5-Difluoroaniline	2.00	1.54	1.64	343	Nonanoic acid/pelargonic	6.65	4.97	2.70
129	1,3-Difluorobenzene	0.53	0.55	1.45	344	1-Nonanol/nonyl alcohol	8.55	7.23	2.25
130	1,4-Difluorobenzene	0.59	0.55	1.45	345	5-Nonanone/dibutyl ketone	1.16	0.97	1.69
131	1,1-Difluoroethane/R152a	0.16	0.19	0.61	346	Octadecane	4.13	3.84	3.84
132	1,1-Difluoroethyl acetate	1.00	0.75	1.12	347	Octane	0.50	0.46	1.31
133	2,2-Difluoroethyl alcohol	2.20	2.43	0.97	348	i-Octane	0.47	0.50	0.49
134	Diheptylamine	2.94	3.79	3.18	349	1-Octanethiol	1.22	1.15	1.21
135	Diisobutylamine	0.70	0.83	0.93	350	2-Octanethiol	1.04	1.02	0.87
136	Diisopentylamine	1.25	1.27	1.16	351	Octanoic acid/caprylic acid	4.75	4.02	2.43
137	Diisopropenyl/diallyl	0.26	0.28	0.61	352	Trans-2-octene	0.47	0.33	1.13
138	Diisopropylamine	0.38	0.40	0.80	353	Octylamine	1.44	1.71	2.07
139	1,2-Dimethoxyethane	0.40	0.38	0.98	354	Paracetaldehyde/paraldehyde	1.04	0.86	1.30
140	N,N-Dimethylacetamide	1.89	1.89	0.81	355	Pentachloroethane	2.18	1.94	0.38
141	Dimethylamine	0.18	0.34	0.79	356	Pentadecane	2.45	2.03	2.78
142	N,N-Dimethylaniline	1.26	1.26	1.79	357	Pentafluoroethane/R125	0.14	0.18	0.49
143	2,6-Dimethylaniline/2,6-xylidine	3.08	2.50	1.19	358	Pentane	0.22	0.24	0.94
144	2,2-Dimethylbutane/neohexane	0.34	0.35	0.60	359	2-Pentanethiol/sec-amyl mercaptan	0.51	0.54	0.63
145	Dimethyl carbonate	0.57	0.51	1.14	360	1-Pentanol/(n-)amyl alcohol	3.42	3.09	1.45
146	N,N-Dimethylethyleneurea	1.88	1.88	1.28	361	2-Pentanol	3.22	2.18	1.16
147	N,N-Dimethylformamide	0.78	0.79	0.97	362	2-Pentanone/methyl propyl ketone	0.46	0.47	1.03
148	2,6-Dimethy-4-heptanone	0.82	0.76	0.94	363	3-Pentanone/diethyl ketone	0.43	0.42	1.10
149	2,7-Dimethyloctane	0.79	0.55	0.91	364	2-Pentene	0.20	0.17	0.82
150	1,4-Dimethyl-5-octyldecahydronaphthalene	14.22	13.11	4.13	365	Penthyl ether	0.97	1.05	1.87
151	1,4-Dimethyl-5-octylnaphthalene	22.02	22.02	2.93	366	Perfluoro-1-isopropoxy hexane	1.49	1.51	0.22
152	4,4-Dimethyl-2-oxazolidinone	87.07	87.06	1.17	367	Phenethyl alcohol/2-phenylethanol	10.30	8.77	2.95
153	2,4-Dimethyl-3-pentanone	0.58	0.60	0.79	368	Phenol	8.08	7.87	2.66
154	2,4-Dimethylsulpholane	8.54	9.42	1.49	369	Phenyl-N-amyl ether	2.05	2.14	3.46
155	1,3-Dimethyl-3,4,5,6-tetrahydro-2(1H)pyrimidinone	2.81	2.88	2.06	370	2-Phenyl-1-propanol	13.90	11.82	2.67
156	1,4-Dioxane	1.14	0.87	3.56	371	Phenyl propyl ketone	2.03	2.40	2.95
157	1,3-Dioxolane	0.58	0.92	2.97	372	Phosphorous trichloride	0.52	0.52	1.44
158	Dipentylamine	1.39	1.62	2.07	373	Phosphoryl chloride	0.99	1.10	2.51
159	Dipheylamine	15.62	9.27	6.52	374	2-Picoline/2-methyl pyridine	0.74	0.85	1.88
160	Dipropylamine	0.50	0.69	1.35	375	Pinacolone/tert-butyl methyl ketone	0.65	0.65	0.60
161	1,1-Di-m-tolylethane	7.33	6.65	2.12	376	Piperidine	1.51	1.10	3.33
162	1,1-Di-p-tolylethane	4.60	6.74	2.12	377	1-Propanamine/propylamine	0.37	0.59	1.21
163	Dodecane	1.34	1.07	2.01	378	2-Propanamine/isopropylamine	0.31	0.36	0.92
164	Ethanethiol/ethyl mercaptane	0.28	0.32	0.63	379	1-Propanethiol/propyl mercaptan	0.38	0.40	0.70
165	Ethanol/ethyl alcohol	1.04	1.63	1.06	380	2-Propanethiol/isopropyl mercaptan	0.35	0.36	0.49
166	Ether/ethyl ether/diethyl ether	0.61	0.29	0.99	381	1-Propanol/n-propyl alcohol	0.86	2.02	1.18
167	2-Ethoxybenzenamine	4.72	5.18	2.45	382	2-Propanol/isopropyl alcohol	1.11	1.45	0.91
168	4-Ethoxybenzenamine	9.41	5.36	2.84	383	Propionaldehyde/prapanal	0.32	0.41	1.12

169	Ethoxybenzene	1.16	1.13	2.50	384	Propionic acid	1.00	1.39	1.43
170	2-Ethoxyethanol	1.77	1.66	1.52	385	Propionic anhydride	1.01	0.96	1.81
171	2-(2-Ethoxyethoxy)ethanol	3.81	3.80	2.20	386	Propionitrile	0.40	0.46	1.10
172	2-[2-(2-Ethoxyethoxy)ethoxy]ethanol	6.82	8.67	3.18	387	Propionyl chloride/prapanyl chloride	0.44	0.53	0.89
173	2-Ethoxytoluene	1.24	1.29	1.70	388	Propoxybenzenephenyl propyl ether	1.39	1.40	2.78
174	3-Ethoxytoluene	1.09	1.17	1.70	389	Propyl acetate	0.53	0.53	1.26
175	4-Ethoxytoluene	1.14	1.17	1.70	390	Propyleneglycol dinitrate	3.38	3.82	5.90
176	Ethyl acetate	0.41	0.43	1.14	391	Propyl ether	0.39	0.45	1.22
177	2-Ethylaniline	3.61	3.56	2.18	392	Propyl formate	0.48	0.49	1.50
178	N-Ethylaniline	1.96	1.92	2.67	393	Propyl oleate	6.32	8.98	6.67
179	Ethyl benzene	0.61	0.69	1.92	394	Propyl sulfone/dipropyl sulfone	6.85	6.85	1.97
180	Ethyl butyrate	0.63	0.55	1.48	395	Propyl sulfoxide/dipropyl sulfoxide	5.99	5.18	1.46
181	Ethyl caprate/ethyl decanoate	2.02	1.96	2.83	396	Propyl thioacetate	0.82	0.87	1.11
182	Ethyl caproate/ethyl hexanoate	0.92	0.84	1.84	397	2-Propyltridecanol	25.36	27.96	3.82
183	Ethyl cyclohexane	0.77	0.80	2.33	398	Pyridine	0.86	0.83	2.77
184	Ethylenediamine	1.45	1.70	1.72	399	Pyrrole	1.19	0.95	2.48
185	Ethylene glycol dinitrate	3.52	3.54	6.64	400	Pyrrolidine	0.69	1.17	2.80
186	Ethyl fluoroacetate	0.86	0.59	1.23	401	Salicylic acid	2.40	2.64	4.04
187	N-Ethyl formamide	2.13	2.13	1.47	402	Spiro(4.5)decane	1.95	2.90	5.30
188	Ethyl formate	0.37	0.39	1.35	403	Spiro(5.5)undecane	2.73	2.73	6.36
189	Eethyl heptanoate/ethyl heptylate	1.12	1.04	2.05	404	Spiro(5.6)dodecane	3.83	2.58	7.49
190	2-Ethyl-1-hexanol	5.82	5.11	1.62	405	Styrene/vinyl benzene	0.68	0.68	2.05
191	2-Ethyl-1-hexene	0.44	0.41	1.01	406	1,1,2,2-Tetrabromoethane	9.17	9.26	2.23
192	Ethyl-lactate	2.26	4.30	1.51	407	1,2,3,4-Tetrahydro-6-butyl-7-hexylnaphtalene	20.96	20.96	6.07
193	Ethyl laurate	2.87	3.00	3.52	408	1,2,3,4-Tetrahydro-7-butyl-1-hexylnaphtalene	15.33	15.33	6.13
194	Ethyl-N-methyl carbamate	2.94	2.94	1.51	409	Tetrahydrofuran	0.46	0.83	2.63
195	Ethyl myristate	4.04	4.59	4.35	410	1-o-Tolyl-1-p-tolylethane	9.80	7.38	2.12
196	3-Ethyl-2-oxazolidinone	2.71	2.88	3.00	411	Tributyl phosphate	3.28	3.87	3.23
197	Ethyl palmitate	5.45	7.02	5.41	412	Tridecane	1.67	1.33	2.23
198	Ethyl propionate	0.49	0.44	1.33	413	Trifluoroacetic acid	0.79	1.09	1.02
199	Ethyl stearate	7.05	10.74	6.68	414	1,1,1-Trifluoro-tert-butyl alcohol	3.07	2.72	0.58
200	Ethyl sulfide	0.41	0.37	1.00	415	1,1,1-Trifluoroethane/R143a	0.11	0.12	0.50
201	Ethyl tert-butyl ether	0.49	0.55	0.65	416	3,3,3-Trifluoro ethanol	1.67	1.67	0.80
202	Ethyltrifluoro acetate	0.40	0.40	0.96	417	1,1,1-Trifluoro isopropyl alcohol	2.35	1.37	0.74
203	Ethyl valerate	0.66	0.68	1.66	418	Trifluoromethylcyclohexane	0.87	0.87	1.75
204	Eugenol/4-allyl-2-methoxyphenol	6.79	8.81	2.33	419	alfa,alpha,alpha-Trifluorotoluene	0.53	0.53	1.44
205	2-Fluoroaniline	2.22	2.93	2.07	420	1,1,2-Trifluoro 1,2,2-trichloroethane	0.64	0.49	0.34
206	3-Fluoroaniline	2.37	2.66	2.07	421	Trihexylamine	3.78	3.78	3.70
207	4-Fluoroaniline	2.55	2.68	2.07	422	Trimethyl phosphite	0.53	0.53	0.56
208	Fluorobenzene	0.54	0.57	1.82	423	Tri-n-propyl ortho phosphate	2.41	2.05	2.33
209	2-Fluoroethanol	1.64	2.11	1.09	424	Tridimethylaminophospine oxide	3.24	3.90	0.68
210	1-Fluorohexane/hexyl fluoride	0.45	0.39	1.09	425	Undecane	1.09	0.87	1.81
211	1-Fluoro-2-nitrobenzene	2.34	2.09	3.42	426	2-Undecanone/methyl nonyl ketone	1.80	1.70	1.96
212	1-Fluoro-3-nitrobenzene	2.11	1.90	3.42	427	Valeronitrile	0.69	0.70	1.36
213	1-Fluoro-4-nitrobenzene	2.41	1.91	3.42	428	m-Xylene/1,3-dimethyl benzene	0.57	0.62	1.05
214	1-Fluoropentane/N-Amyl fluoride	0.35	0.31	0.97	429	o-Xylene/1,2-dimethyl benzene	0.74	0.68	1.05
215	2-Fluorophenetole	1.32	1.21	2.39	430	p-Xylene/1,4-dimethyl benzene	0.59	0.62	1.05

Table B.2. Eight experimental data points that cannot be represented by the M&G groups. They were added to the list of compounds of Table B1 for the atom parameter regression of the M&G viscosity model. Experimental data and CI estimations.

n° name		***exp***	**CI-*est***
1	1.3-bis(1-Isocyanato-1-methylethyl)benzene	7.77	2.96
2	Carbon disulfide	0.35	0.79
3	2,6-Dimethylpyridine-N-oxide	5.76	1.52
4	Ethylene sulfite	1.99	3.36
5	1-(Isocyanato-1 methylethyl)-4-(1-methylethenyl)benzene	3.75	1.95
6	Phenyl isothiocyanate/isothiocyanatobenzene	1.39	2.61
7	1,3-Propane sulfone	9.16	2.22
8	Thiophosphoryl chloride/phosphorous sulfochloride	1.00	2.19

C

M&G GC$^+$:

groups, atoms and contributions

Tables C.1-C.3 list the first-, second- and third-order groups (C_i, D_j, E_k) and their contributions for surface tension and viscosity predictions. For each group, a compound example is given, together with the frequency (N_i, M_j, O_k) of the groups in the compound (in brackets).

Table C.1. First-order group and their contributions for surface tension and viscosity.

group	example (N_i)	$C_i - \sigma$	$C_i - \eta$
CH_3	Acetaldehyde (1)	8.0328	-1.0278
CH_2	Allyl alcohol (1)	0.6213	0.2125
CH	2-Butanol (1)	-7.7843	1.3180
C	2-Methyl-2-Propanol (1)	-16.3927	2.8147
CH_2=CH	3-Chloropropene (1)	9.7658	-0.9970
CH=CH	Cis-Pentene (1)	0.7405	0.0739
CH_2=C	Diisopropenyl (2)	/	0.3612
CH=C	2-Methyl-2-Butene (1)	-6.9485	1.4719
CH#C	1-Hexyne (1)	/	-0.6725
aCH	Benzene (6)	4.8383	-0.0901
aC	Naphtalene (2)	-0.2411	1.0776
aC	1,2,3,4-Tetrahydro-Naphthalene (2)	-1.5409	0.7050
aC	Diphenylamine (1)	-7.7616	1.3376
aN	Methyl pyridine (1)	10.5336	0.2666
aC-CH_3	1,2,3-Trimethylbenzene (2)	4.3657	-0.0576
aC-CH_2	Ribenzylamine (3)	-4.2980	1.1035
aC-CH	p-Cymene (1)	-13.2933	2.4299
aC-C	Tert-Butylbenzene (1)	-20.7377	2.0188
aC-CH=CH_2	Styrene (1)	/	0.0667
aC-C=CH_2	1-(Isocyanato-1-Methylethyl)-3-(1-Methylethenyl)benzene (1)	/	1.3080
OH	Cyclohexanol (1)	16.0184	1.3057
aC-OH	2-Chlorophenol (1)	11.6969	2.5133
COOH	Acetic Acid (1)	16.9300	1.1430
aC-COOH	Salicylic Acid (1)	/	-1.2767
CH_3CO	Acetic Anhydride (1)	15.3216	-0.1881
CH_2CO	5-Nonanone	7.2453	0.9647

CHCO	2,4-Dimethyl-3-Pentanone (1)	/	2.2504
aC-CO	Benzoylbromide (1)	3.5743	1.9268
CHO	Formamide (1)	14.5740	-0.076
aC-CHO	Benzaldehyde (1)	16.1730	/
CH_3COO	Ethyl Acetate (1)	14.8168	-0.0358
CH_2COO	Ethyl Butanoate (1)	7.1186	1.0292
CHCOO	Ethyl Isobutyrate (1)	-1.1233	1.9966
CCOO	Ethyltrifluoro Acetate (1)	/	2.1046
HCOO	Isobutyl Formate (1)	14.7142	-0.1208
aC-COO	Methyl Benzoate (1)	1.0627	1.9603
aC-OOCH	Phenylsalicylate (1)	18.6085	/
COO	Diethyl Maleate (1)	7.2230	/
CH_3O	Butyl Methyl Ether (1)	9.7810	-0.6902
CH_2O	Dibutyl Ether (1)	2.3925	0.6134
CH-O	N,N-Dimethylformamide (1)	-7.0770	3.6344
aC-O	Anisole (1)	-2.3913	1.3912
CH_2NH_2	Butylamine (1)	16.3950	0.2902
$CHNH_2$	Sec-Butylamine (1)	4.4131	1.0108
CNH_2	Tert-Butylamine (1)	-7.2285	/
CH_3NH	Dimethylamine (1)	18.3072	-0.0637
CH_2NH	Dipropylamine (1)	4.5002	1.0512
CHNH	Diisopropylamine (1)	-5.2070	1.8378
CH_3N	N,N-Dimethylformamide (1)	-2.6557	0.8715
CH_2N	Tributylamine (1)	-4.8682	1.4376
aC-NH_2	Aniline (1)	14.7535	1.4614
aC-NH	N-Ethylaniline (1)	3.3787	1.9164
aC-N	N-N-Dimethylaniline	-6.1185	2.7340
NH_2	Cyclohexylamine (1)	17.5324	0.0733
C=N	1-(Isocyanato-1-Methylethyl)-3-(1-ethylethenyl)benzene (1)	/	1.3080
CH_2CN	Hexanenitrile (2)	21.4598	0.2417
aC-CN	p-Tolunitrile (1)	13.9398	0.7746
CN	Ethyl Thiocyanate (1)	20.5959	-0.2253
$CH2NO_2$	Cychloropentanone (1)	22.7715	/
$CHNO_2$	2-Nitropropane (1)	13.2243	/
aC-NO_2	o-Chloronitrobenzene (1)	16.8215	1.1257
NO_2	Nitromethane (1)	28.4972	0.2776
ONO_2	Diethylen Glycol Dinitrate (2)	/	0.4193
$HCONHCH_2$	N-Ethyl Formamide (1)	/	1.7859
$CONH_2$	2-Phenylacetamide (1)	26.2762	/
$CONHCH_3$	Ethyl-N-Methyl Carbamate (1)	/	1.4921
$CONHCH_2$	N-2-Methoxyethyl acetamide (1)	/	3.8174
$CON(CH_3)_2$	N,N-Dimethylacetamide (1)	/	1.6644
aC-$CONH_2$	Benzamide (1)	21.3085	/
aC-NH(CO)H	Formanilide (1)	17.9085	1.9704
aC-NHCO	Acetanilide (1)	11.6757	/
NHCON	N,N-Dimethylethyleneurea (1)	/	3.5019
CH_2Cl	1-Chlorobutane (1)	14.9134	-0.1898
CHCl	2-Chlorobutane (1)	4.4604	0.8439
CCl	2-Methyl-2-Chloropropane (1)	/	2.3105

$CHCl_2$	1,1-Dichloroethane (1)	17.4043	0.2334
CCl_3	Pentachloroethane (1)	16.2169	0.4290
CCl_2F	Dichlorodifluoromethane (1)	/	-0.3579
CH_2F	1-Fluorohexane (1)	9.9927	-0.7707
CHF_2	1-Bromo-2,2-Difluoroethane (1)	/	-0.6292
CF_2	Perfluoro-1-Isopropoxy Hexane (5)	/	-0.0227
CF_3	Trifluoracetic Acid (1)	-3.4000	-1.0660
HCClF	Chlorodifluoromethane (1)	/	-0.7687
$CClF_2$	1-Chlorodifluoroethane (1)	/	-0.3643
aC-Cl	1-Chloronaphtalene (1)	8.4233	0.1867
aC-F	m-Fluorotoluene (1)	3.9769	-0.1196
aC-I	p-Iodo-2-Methylpropane (1)	13.7998	0.8645
aC-Br	1-Bromonaphtalene (1)	10.6992	0.5262
I	1-Iodo-2-Methylpropane (1)	18.9248	0.2829
Br	1-Bromodecane (1)	17.0377	-0.1025
F	Chlorodifluoromethane (1)	/	-0.7351
Cl	Dichloromehane (1)	11.6721	-0.5777
OCH_2CH_2OH	2-Butoxyethanol (1)	19.5821	1.3234
CH_2SH	Ethanethiol (1)	15.5059	-0.1111
CHSH	2-Propanethiol (1)	5.2643	1.0169
CSH	2-Methyl-2-Butanethiol (1)	/	2.3456
SH	Methanethiol (1)	15.8672	/
CH_3S	Dimethyl disulfide (2)	16.4255	-0.1698
CH_2S	Dipenthyl alculat (1)	6.8133	0.8614
aC-S-	Ethyl phenil alculat (1)	5.5650	/
SO	Dimethyl sulfoxide (1)	26.8543	2.8511
SO_2	Sulfuryl Chloride (1)	5.4359	3.1292
SO_4	Diethylsulfate (1)	15.7918	/
$aC\text{-}SO_2$	p-Toluenesulfonylchloride (1)	4.8090	/
P	Phosphorus Trichloride (1)	-7.0362	1.0810
PO_3	Trimethyl Phosphite (1)	/	2.4516
PO_4	Triethyl phosphate (1)	3.5377	2.5240
CO_3	Diethyl Carbonate (!)	9.6584	1.3754
C_2H_3O	Epichloridrin (1)	21.4466	/
CH_2(cyc)	Pyrrolidine (4)	4.2246	-0.0577
CH(cyc)	2-Methylcyclohexanone (1)	-5.7642	0.9455
C(cyc)	4,4-Dimethyl-2-Oxazolidinone (1)	/	1.5824
CH=CH(cyc)	Cyclohexene (1)	11.7426	-0.2162
CH=C(cyc)	Forfural (1)	9.9907	0.6127
CH_2=C(cyc)	4,4-Dimethyl-2-Oxazolidinone (1)	/	4.5686
NH(cyc)	Piperidine (1)	11.2310	0.3855
N(cyc)	1-Methyl-2-Pyrrolidone (1)	6.5739	1.4649
O(cyc)	1,4-Dioxane (2)	6.7827	0.0434
CO(cyc)	Cyclohexanone (1)	12.9294	0.5813
S(cyc)	Thiophene (1)	7.1948	/
SO_2(cyc)	2,4-Dimethylsulpholane (1)	/	2.5441
> NH	Dibenzylamine (1)	0.8130	/
-O-	Benzyl Ether (1)	/	0.1796
Si	Tetrachlorosilane (1)	-28.3983	/

P=O	Phosphoryl Chloride (1)	-2.9862	1.8303

Table C.2. Second-order group and their contributions for surface tension and viscosity.

group	example (M_j)	$D_j - \sigma$	$D_j - \eta$
$(CH3)_2CH$	2-Nitropropane (1)	-0.1221	0.0142
$(CH3)_3C$	2,2-Dimethylpentane (1)	0.3379	0.0773
$CH(CH_3)CH(CH_3)$	2,3-Dimethylbutane (1)	1.3494	0.4075
$CH(CH_3)C(CH_3)_2$	2,2,3-Trimethylpentane (1)	3.2537	/
$CHn=CH_m-CH_p=CH_k$ (k,m,n,p in 0..2)	1,3-Butadiene (1)	/	-0.0191
$CH_3-CH_m=CH_n$ (m,n in 0..2)	Ethyl Crotonate (1)	0.1393	0.0380
$CH_2-CH_m=CH_n$ (m,n in 0..2)	Allyl Acetate (1)	-0.1933	-0.0198
CHCHO	Chloral (1)	/	-0.1430
CH_3COCH_2	2,4-Hexadione	-0.4544	0.0454
CH3COCH	3,3-Dimethyl-2-Butanone	/	-0.0500
CHCOOH	Trifluoroacetic Acid (1)	-0.2446	0.0055
CH_3COOCH	Tert-Butylacetate (1)	-0.9320	-0.2373
CO-O-CO	Acetic Anhydride (1)	1.7916	0.0259
CHOH	3-methyl-2-Pentanol (1)	-2.3870	-0.2116
COH	2-Methyl-2-Butanol (1)	-2.6119	0.0033
$OH-CH_n-COO$ (n in 0..2)	Ethyl Lactate (1)	-3.2820	/
$CH_m(OH)CH_n(OH)$ (m,n in 0..2)	1,2-Ethnediol (1)	14.7106	/
$CH_m(OH)CH_n(NH_p)$ (m,n,p in 0..2)	Ethanolamine (1)	15.2853	0.6128
$CH_m(NH_2)CHn(NH_2)$ (m,n in 0..2)	1,2 Propanediamine (1)	/	-0.0484
$CH_m(NH)CH_n(NH_2)$ (m,n in 1..2)	2-((2-Aminoethyl)amino)Ethanol (1)	/	0.7947
$NC-CH_n-CH_m-CN$ (n, m in 1..2)	Succinonitrile (1)	7.6805	/
$NC-CH_n-COO$ (n in 1..2)	Methylcycanoacetate (1)	0.3737	/
$COCH_nCOO$ (n in 1..2)	Ethyl Acetoacetate (1)	1.9209	/
$CH_m=CH_n-Br$ (m,n in 0..2)	1,2-Dibromoethylene (1)	/	0.0086
$CH_m=CH_n-Cl$ (m,n in 0..2)	Trans-1,2-Dichloroethylene (2)	/	0.0528
$CH_n=CH_m-COO-CH_p$ (m,n,p in 0..3)	Ethyl Fumarate (2)	0.2242	/
$aC-CH_n-X$ (n in 1..2) X: Halogen	Benzyl chloride (1)	/	0.1365
$aC-CH_n-NH_m$ (n in 1..2; m in 0..2)	Benzylamine (1)	0.3748	-0.2742
$aC-CH_n-OH$ (n in 1..2)	Benzyl Alcohol (1)	-1.1119	-0.3156
$aC-CH_n-CN$ (n in 1..2)	Phenylacetonitrile (1)	1.2106	0.2099
$aC-CH_n-CONH2$ (n in 1..2)	Phenylacetamine (1)	-1.8697	/
$aC-CH_n-OOC$ (n in 1..2)	Benzyl Benzoate (1)	-2.3277	-0.0988
$aC-CH(CH_3)_2$	Isopropylbenzene (1)	-0.0083	-0.2512
$aC-C(CH_3)_3$	Tert-Butylbenzene (1)	-0.1666	/
$Chcyc-CH_3$	Methylcyclohexane (1)	-0.3931	-0.0108
$Chcyc-CH_2$	Epichlorohydrine (1)	0.5685	-0.0620
Chcyc-C	Trifluoromethylcyclohexane (1)	/	0.2663
Chcyc-Cl	Cyclohexyl Chloride (1)	/	0.2714
Chcyc-OH	Cycloheptanol (1)	1.3649	0.8711
$Chcyc-NH_2$	Cyclohexylamine (1)	-1.6714	-0.1078
$> Ncyc-CH_3$	N-Methyl-2-Pyrrolidone (1)	/	-0.1127
$> Ncyc-CH_2$	3-Butyl-2-Oxazolidone (1)	/	-0.1015
AROMRINGs1s2	1-Bromo-2-Methyl Benzene (1)	0.5195	0.0930
AROMRINGs1s3	1-Bromo-3-Methyl Benzene (1)	0.5487	-0.0041
AROMRINGs1s4	1-Bromo-4-Methyl Benzene (1)	-0.0303	0.0027

AROMRINGs1s2s3	1,2,3-Trimethyl Benzene (1)	0.688	-0.1610
AROMRINGs1s2s4	1,2,4-Trimethyl Benzene (1)	-0.4597	-0.5174
AROMRINGs1s3s5	1,2,5-Trimethyl Benzene (1)	-1.4526	/
PYRIDINEs2	2-Methyl Pyridine (1)	-1.2525	-0.0069
PYRIDINEs3	3-Methyl Pyridine (1)	0.2475	/
PYRIDINEs4	4-Methyl Pyridine (1)	0.3332	/

Table C.3. Third-order group and their contributions for surface tension and viscosity.

group	**example (O_k)**	E_k - σ	$E_k - \eta$
NC-(CHn)m-CN (m > 2)	Hexanitrile (1)	1.2879	/
aC-(CHn=CHm)cyc (fused rings) (n,m in 0..1)	Indene (1)	/	0.3471
aC-aC (different rings)	Biphenyl (1)	6.3401	/
aC-Chncyc (fused rings) (n in 0..1)	1,2,3,4-Tetrahydronaphtalene (2)	-0.4689	0.0417
CH multiring	Butyl Formate (2)	/	-0.0929
aC-CHm-aC (different rings) (m in 0..2)	1,1-Di-p-Tolylethane (1)	/	-0.0012
aC-CO-aC (different rings)	Benzophenone (1)	/	0.0833
aC-CHm-CO-aC (different rings) (m in 0..2)	Benzophenone (1)	/	0.0833
aC-NH-aC (different rings)	Diphenylamine (1)	-1.2001	-0.1258
aC-O-aC (different rings)	Diphenyl Ether (1)	-11.4802	/
AROM.FUSED[2]	1-Methylnaphtalene (1)	0.9379	-0.6857
AROM.FUSED[2]s1	1-Chloronaphtalene (1)	-1.1199	0.0759
AROM.FUSED[2]s2	1,2,3,4-Tetrahydro-7-Butyl-1-Hexylnaphtalene (1)	/	0.2608
AROM.FUSED[2]s2s3	2-Butyl-3-Hexylnaphthalene (1)	/	0.5339
AROM.FUSED[2]s1s4	1,4-Dimethyl-5-Octylnaphthalene (1)	/	0.0759
PYRIDINE.FUSED[2]	Quinoline (1)	-2.2675	/

D

M&G GC^{+}:

application examples

Examples have been developed to illustrate the application of the M&G method and its capability to predict and distinguish between some isomers. For each example, the experimental value (if available) is also reported together with the Relative Deviation (*RD*). The examples are divided in terms of: application of the GC method, §D.1 (all contributions for all the groups that represent the molecule are available) examples of the application of the GC-CI method, §D.2 (some contributions for one or more groups that represent the molecule are not available).

D.1 Application of GC-based method

Example 1. (Table D.1). In this example, the compound under evaluation is quinoline that belongs to the nitrate hydrocarbons family and has contributions of first- and third-order groups. The experimental value of surface tension at 298 K is 42.59 mN/m, while for viscosity no experimental data are available.

Example 2. (Table D.2). In this example, the compound under evaluation is 2-hexanone, that belongs to esters family, and has contributions of first- and second-order. The experimental value of surface tension at 298 K is 25.45 mN/m, while for viscosity no experimental data are available.

Examples 3a-3b. (Table D.3). In these examples the isomers o-nitrotoluene and m-nitrotoluene, which belong to the nitrate hydrocarbons family, are considered. They have contributions of first- and second-order. The experimental value of surface tension for o-nitrotoluene at 298 K is 41.2 mN/m, while for viscosity at 300 K is 2.09 mPa·s. The experimental value of surface tension for m-nitrotoluene at 298 K is 40.8 mN/m, while for viscosity is 2.03 mPa·s.

Table D.1. Example 1. Predictions of surface tension (298 K) and viscosity (300 K) for quinoline.

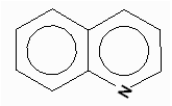
1) Quinoline Cas: 91-22-5 Molecular Formula: C_9H_7N			
1st order group	**occurrences**	**contribution**	
		σ	***η***
aCH	7	4.8383	-0.0901
aC	2	-0.2411	1.0776
aN	1	10.5336	0.2666
	tot =	43.92	1.79
3rd order group	**occurrences**	**contribution**	
		σ	***η***
AROM.FUSED[2]	1	0.9379	-0.6857
PYRIDINE.FUSED[2]	1	-2.2675	0.0000
	tot =	-1.33	-0.69
σ = 43.92 – 1.33 = 42.59 mN/m		*RD* = 0.00%	
η = exp(1.79 – 0.69) = 3.00 mPa·s		*RD* = /	

Table D.2. Example 2. Predictions of surface tension (298 K) and viscosity (300 K) for 2-hexanone.

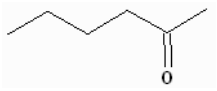
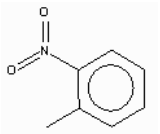
2) 2-Hexanone Cas: 591-78-6 Molecular Formula: $C_4H_9COCH_3$			
1st order group	**occurrences**	**contribution**	
		σ	***η***
CH_3	1	8.0328	-1.0278
CH_2	3	0.6213	0.2125
CH_3CO	1	15.3216	-0.1881
	tot =	25.22	-0.58
2nd order group	**occurrences**	**contribution**	
		σ	***η***
CH_3COCH_2	1	-0.4544	0.0454
	tot =	-0.45	0.05
σ = 25.22 – 0.45 = 24.76 mN/m		*RD* = 2.71%	
η = exp(-0.58 + 0.05) = 0.587 mPa·s		*RD* = /	

Table D.3. Example 3a: predictions of surface tension (298 K) and viscosity (300 K) for o-nitrotoluene. Example 3b: predictions of surface tension (298 K) and viscosity (300 K) for m-nitrotoluene.

3a) o-Nitrotoluene

Cas: 88-72-2

Molecular Formula: $C_7H_7NO_2$



1st order group	occurrences	contribution	
		σ	η
aCH	4	4.8383	-0.0901
$aCCH_3$	1	4.3657	-0.0576
$aCNO_2$	1	16.8215	1.1257
	tot =	40.54	0.71

2nd order group	occurrences	contribution	
		σ	η
AROMRINGs1s2	1	0.5195	0.093
	tot =	0.52	0.09

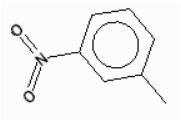
σ = 40.54 + 0.52 = 41.06 mN/m *RD* = 0.34%

η = exp(0.71 + 0.09) = 2.23 mPa·s *RD* = 6.70%

3a) m-Nitrotoluene

Cas: 99-08-1

Molecular Formula: $C_7H_7NO_2$



1st order group	occurrences	contribution	
		σ	η
aCH	4	4.8383	-0.0901
$aCCH_3$	1	4.3657	-0.0576
$aCNO_2$	1	16.8215	1.1257
	tot =	40.54	0.71

2nd order group	occurrences	contribution	
		σ	η
AROMRINGs1s3	1	0.5487	-0.0041
	tot =	0.55	0.00

σ = 40.54 + 0.55 = 41.09 mN/m *RD* = 0.71%

η = exp(0.71 + 0.00) = 2.02 mPa·s *RD* = 0.49%

D.2 Application of GC-CI method (GC^+)

Example 4. (Table D.4). Estimation of the surface tension at 298 K and viscosity at 300 K for diphenylacetylene. It has contributions only of first-order, and the contribution for the group aC-C#C is not available. Experimental values are not available.

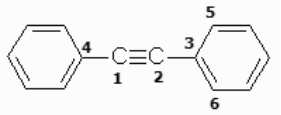
Example 5. (Table D.5). Estimation the surface tension of 2-hexanone at 298 K and viscosity at 300 K (assuming for purposes of illustration that group $COCH_3$ is missing). It has contributions of fisrt- and second-order, and the contribution for the group CH_3CO is not available. Experimental value of surface tension at 298 K is 25.45 mN/m, while for viscosity no experimental data are available.

Table D.4. Example 4. Predictions of surface tension and viscosity (300 K) for diphenylacetylene.

4) Diphenylacetylene

Cas: 501-65-5

Molecular Formula: $C_{14}H_{10}$



Marrero & Gani groups contribution

1st order group	**occurrences**	**contribution** σ	**contribution** η
aC	1	-7.7616	1.0776
aCH	10	4.8383	-0.0901
	tot=	40.62	0.18

2nd order group	**occurrences**	**contribution** σ	**contribution** η
No contribution			

3rd order group	**occurrences**	**contribution** σ	**contribution** η
No contribution			

Atom Connectivity Index contribution Missing GC: aC-C#C

atom	δ	$1/\sqrt{\delta}$
C(1)	4	0.50
C(2)	4	0.50
C(3)	14	0.27
C(4)	14	0.27
C(5)	13	0.28
C(6)	13	0.28

$^{v}\chi^{0} = 0.50 + 0.50 + 0.27 = 1.27$

bond	β	$1/\sqrt{\beta}$
C-C (1-2)	$4 \cdot 4 = 16$	0.25
C-C (2-3)	$4 \cdot 14 = 56$	0.13
C-C (1-4)	$4 \cdot 14 = 56$	0.13
C-C (3-5)	$14 \cdot 13 = 182$	
C-C (3-6)	$14 \cdot 13 = 182$	0.07

$^{v}\chi^{1} = 0.25 + 0.13 + 0.5 \cdot 0.13 + 0.5 \cdot 0.07 + 0.5 \cdot 0.07 = 0.52$

atom	**occurrences**	**contribution** σ	**contribution** η
C	3	2.21	0.31
	tot =	6.63	0.93

$F(\sigma^*) = 6.63 + 1.13 \cdot 1.27 - 7.34 \cdot 0.52 + 0 = 4.25$

$F(\eta^*) = 0.93 - 1.01 \cdot 1.27 + 0.87 \cdot 0.52 + 0 = 0.10$

combined GC and CI methods

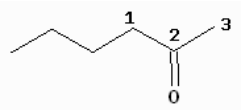
$\sigma = 40.62 + 4.25 = 44.87$ mN/m	*RD* = /
$\eta = \exp(0.18 + 0.10) = 1.32$ mPa·s	*RD* = /

Table D.5. Example 5. Predictions of surface tension and viscosity (300 K) for 2-hexanone, assuming that the group CH_3CO is missing.

5) 2-Hexanone

Cas: 591-78-6

Molecular Formula: $C_4H_9COCH_3$



Marrero & Gani Groups contribution				**Atom Connectivity Index contribution** Missing GC: CH_3CO					
1st order group	**occurrences**	**contribution** σ	η	**atom**	δ	$1/\sqrt{\delta}$	**bond**	β	$1/\sqrt{\beta}$
CH_3	1	8.0328	-1.0278	C(1)	2	0.71	C-C (1-2)	$2 \cdot 4 = 8$	0.35
CH_2	3	0.6213	0.2125	C(2)	4	0.50	C-C (2-3)	$4 \cdot 1 = 4$	
	tot =	9.90	-0.39	C(3)	1	1.00	C(2)-O	$4 \cdot 6 = 24$	0.20
2nd order group	**occurrences**	**contribution** σ	η	O	6	0.41			
				$^{v}\chi^{0} = 0.50 + 1.00 + 0.41 = 1.91$			$^{v}\chi^{1} = 0.5 \cdot 0.35 + 0.50 + 0.20 = 0.88$		
CH_3COCH_2	1	-0.4544	0.0454	**atom**		**occurrences**		**contribution** σ	η
	tot =	-0.45	0.05	C		2		2.21	0.31
3rd order group	**occurrences**	**contribution** σ	η	O		1		6.68	0.49
No contribution					tot =			11.1	1.11

$F(\sigma^*) = 11.1 + 1.13 \cdot 1.91 - 7.34 \cdot 0.88 + 0 = 7.60$

$F(\eta^*) = 1.11 - 1.01 \cdot 1.91 + 0.87 \cdot 0.88 + 0 = -0.77$

Combined GC and CI methods

$\sigma = 9.90 - 0.45 + 7.60 = 17.05$ mN/m — *RD* = 31.10%

$\eta = \exp(-0.39 + 0.05 + 1.11) = 2.16$ mPa· — *RD* = /

E

Viscosity M&G GC$^+$: predictive ability

The predictive ability of the viscosity model was tested on 101 data points not included in the parameter regression. These extra data were collected from the following sources: Van Velzen *et al.*, 1972; Bleazard *et al.*, 1996; Ralph *et al.*, 1992; Katritzky *et al.*, 2000; Miller, 1996; Weng and Chen, 2004.
Table E.1 shows viscosity experimental data and predictions for the 101 compounds employed in testing the predictive ability of the M&G GC$^+$ model.

Table E.1. 101 data points employed for the M&G GC$^+$ viscosity model testing. Experimental values and model estimations.

n^o	name	*exp*	M&G GC$^+$ *est*	n^o	name	*exp*	M&G GC$^+$ *est*
1	2.3-Dimethylbutane	0.35	0.35	52	1-Phenyloctane	4.62	5.76
2	2-Methylhexane	0.36	0.33	53	1-Phenyldodecane	7.32	8.80
3	1-Hexene	0.24	0.24	54	1-Phenylpentaecane	7.32	10.89
4	Propylbenzene	0.78	0.85	55	Hexanoic acid	2.73	2.63
5	1-Methylnaphtalene	2.91	4.34	56	Oleic acid	29.08	22.75
6	Methanol	0.54	1.32	57	1,1,2-Trichlorotrifluoroethane	0.65	0.49
7	2-Methyl-2-butanol	3.53	3.50	58	2-Hydroxybenzaldheyde	2.47	9.45
8	Ethyl propyl ether	0.29	0.36	59	o-Phenetidine	4.69	5.87
9	Dibutyl ether	0.63	0.68	60	p-Phenetidine	9.51	5.36
10	Aacetophenone	1.65	1.57	61	1-Octene	0.43	0.37
11	Ethylbenzoate	1.95	2.00	62	Tetralin	1.91	2.27
12	m-Toluidine	3.03	2.83	63	1-Octanol	7.07	5.84
13	Nitromethane	0.61	0.47	64	Diethyl ether	0.22	0.68
14	Thiophene	0.61	0.65	65	2-Butanone	0.40	0.38
15	Chlorobenzene	0.73	0.77	66	1,2-Ethanediamine	1.31	1.70
16	Morpholine	1.92	1.22	67	3-Methyl-1-butanol	3.47	2.74
17	Tetradecane	1.93	1.64	68	Methoxybenzene	1.28	0.92
18	1-Nonene	0.66	0.46	69	Propylformate	0.49	0.49
19	1,2-Propanediol	31.12	22.51	70	Pentyl acetate	0.88	0.81
20	1,3-Propanediol	37.32	25.76	71	Methyl benzoate	1.78	1.62
21	Triethylene glycol	35.09	21.58	72	Dibutyl-o-phtalate	15.19	17.68
22	Methylphenylamine	2.00	1.55	73	Triethylamine	0.37	0.29
23	N-Methylpropionamide	5.01	1.97	74	o-Toluidine	3.70	3.12

24	1-Bromobutane	0.59	0.61
25	Tetrachloroethylene	0.83	0.12
26	1,5-Hexadiene	0.26	0.20
27	1-Dodecene	1.17	0.88
28	1-Tridecene	1.45	1.08
29	1-Hexadecene	2.60	2.05
30	1-Heptadecene	3.10	2.54
31	1-Octadecene	3.67	3.14
32	Butylcyclopentane	1.18	1.30
33	Propylcyclohexane	0.91	0.99
34	1,3,5-Trimethylcyclohexane	0.65	0.64
35	Butylcyclohexane	0.80	1.23
36	N-Amylcyclohexane	1.52	1.52
37	N-Hexylcyclohexane	1.93	1.88
38	N-Heptylcyclohexane	2.42	2.32
39	N-Nonylcyclopentane	2.60	3.76
40	N-Decylcyclopentane	3.02	4.65
41	N-Undecylcyclopentane	3.75	5.75
42	N-Undecylcyclohexane	5.22	5.43
43	N-Tridecylcyclopentane	5.06	8.80
44	N-Tetradecylcyclopentane	5.96	10.89
45	N-Pentadecylcylcyclopentane	6.89	13.46
46	N-Hexadecylcylcyclopentane	7.91	16.65
47	1-Methyl-4-ethylbenzene	0.64	0.71
48	Amylbenzene	1.24	1.30
49	1-Phenyloctane	2.34	2.46
50	1-Phenylnonane	2.86	3.04
51	1-Phenylundecane	3.99	4.65
75	Nitroethane	0.65	0.36
76	3-Bromoaniline	5.46	5.07
77	1-Undecene	0.93	0.71
78	1-Tetradecene	1.77	1.34
79	1-Pentadecene	2.17	1.66
80	Ethylcyclopentane	0.51	0.85
81	Propylcyclopentane	0.62	1.05
82	N-Amylcyclopentane	1.03	1.61
83	N-Hexylcyclopentane	1.34	1.99
84	N-Heptylcyclopentane	1.67	2.46
85	N-Octylcyclopentane	2.06	3.04
86	N-Octylcyclohexane	1.00	2.87
87	N-Nonylcyclohexane	3.66	3.55
88	N-Decylcyclohexane	4.40	4.39
89	N-Dodecylcyclopentane	4.32	7.12
90	N-Dodecylcyclohexane	6.15	6.72
91	N-Tridecylcylcyclohexane	7.27	8.31
92	1-Phenylheptane	1.97	1.99
93	1-Phenyldecane	3.39	3.76
94	1-Phenyltridecane	5.39	7.12
95	Pentanoic acid	1.99	2.12
96	Chloroform	0.51	0.71
97	Benzophenone	12.46	10.63
98	N,N-Dimethylethanolamine	3.07	4.83
99	N,N-Diethylethanolamine	3.61	3.76
100	N-Methyldiethanolamine	61.71	76.16
101	N-Ethyldiethanolamine	51.26	48.00

F

The virtual PPD-lab: formulations work-flow

A detailed description of the framework implemented in the virtual PPD-lab for the design and verification of formulated products is given in this appendix. This framework is based on the work-flow of stage 1 of the methodology presented in Chapter 5. The virtual PPD-lab allows virtual experimentations related to finding new formulations (use of the design option) and/or verifying/improving an existing design (use of the verification option).

F.1 Task 1: problem definition

This task corresponds to tasks S1-D1 and S1-V1 of the methodology for the design and verification of formulated products (see Chapter 5, Fig. 5.1).

Step 1.1: problem description. In this step the user selects which product to design or to verify. There are different options to select from, as shown in Fig. F.1.
For a set of products, the work-flow has already been implemented. These products correspond to the case studies developed in this work: design of an insect repellent lotion, sunscreen lotion, paint formulation (see Chapter 6), verification of a hair spray formulation (see Chapter 7). If the user wants to design another type of product or a new product, the option 'other' has to be selected. The user can also input information found in the literature about their specific product and the corresponding references.

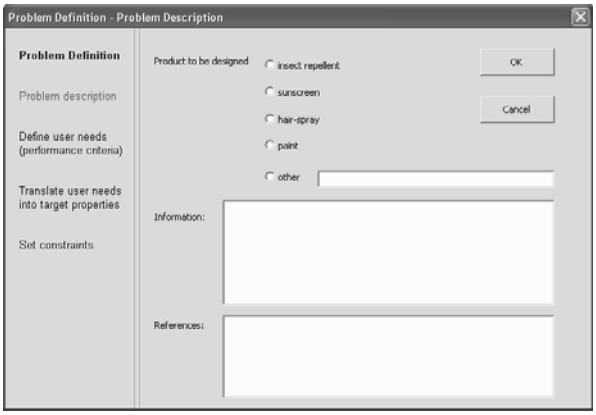


Figure F.1. The problem definition main menu, where the specific problem to solve has to be selected.

Step 1.2: define user needs. In this step, the user selects the product performance criteria (or user needs). The knowledge base has been added to the virtual PPD-lab to guide the user, as shown in Fig. F.2. For instance, if the user selects the sunscreen lotion from the dialogue box of Fig. F.1, all the performance criteria discussed in Chapter 6 for a sunscreen lotion are shown, and the user has only to select the criteria he/she wants to include for the product under development. In addition, new performance criteria can be added in the space reserved for 'others', at the bottom of the dialogue box.

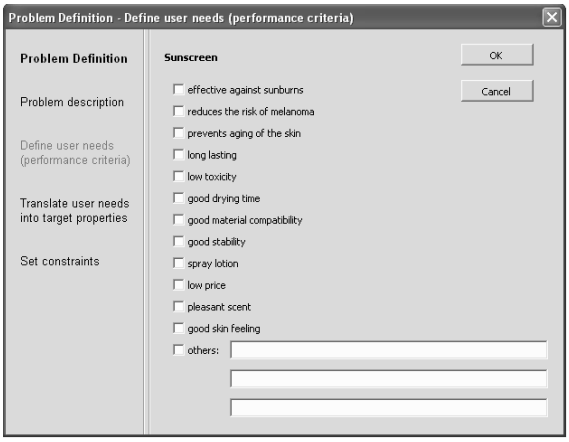


Figure F.2. The performance criteria (user needs) dialogue box for one of the products in the knowledge base (sunscreen lotion).

If in step 1.1 (see Fig. F.1) the user had selected ‘other’, for example ‘detergent’, he/she can now (in step 1.2) introduce his/her user defined needs as shown in Fig. F.3 (toxicity, drying time, water rinsability).

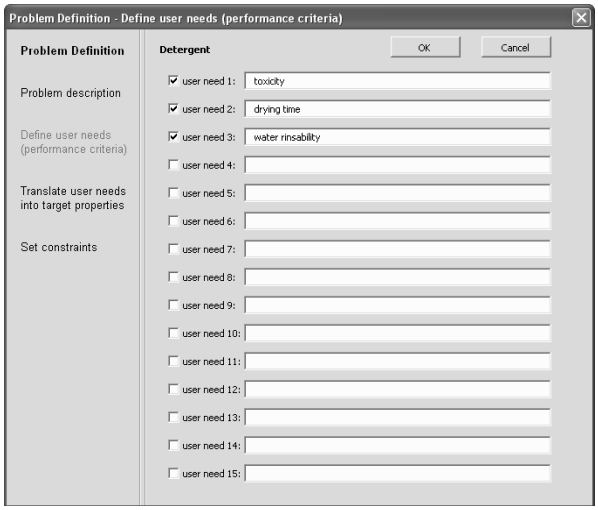


Figure F.3. The performance criteria (user needs) dialogue box for a user defined product, a detergent in this case.

Step 1.3: translate user needs. In this step, the selected performance criteria (from step 1.2) are translated into physicochemical properties (target properties) as shown in the dialogue box of Fig. F.4. Also here, the developed knowledge base is used by the virtual PPD-lab for solving the case studies. That is, in the knowledge base (Table 3.1, Chapter 3), the target properties are already assigned to the corresponding user needs. The user is allowed to change the properties if necessary, or assign new properties to the same attributes, as shown in Fig. F.4.

In the case the option ‘other’ had been selected for the product type in the dialogue box of Fig. F.1 (i.e., ‘detergent’), the user will now need to assign at least one target property to each of the user needs (attributes) that were added in the dialogue box of Fig. F.3.

The ‘Help’ in the top right corner of the dialogue box of Fig. F.4 gives the information needed by the user about the translation process (mainly, the meaning of the symbols used to represent the physicochemical properties).

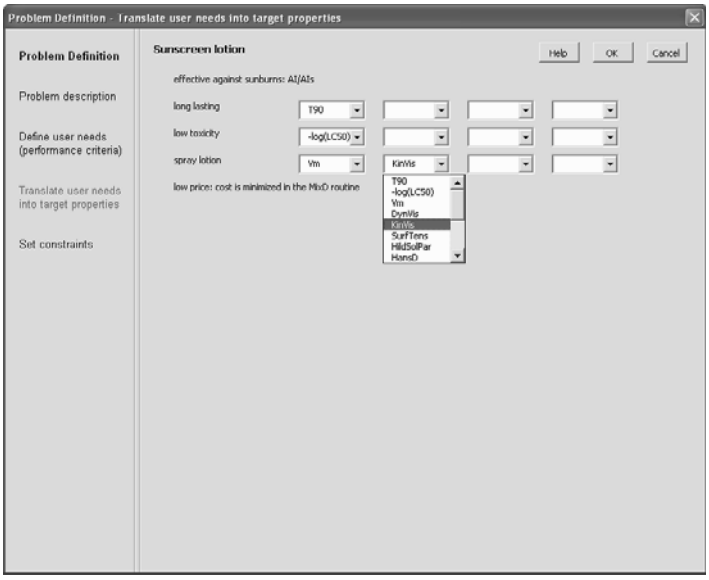


Figure F.4. The dialogue box where the user needs are translated into target properties.

Step 1.4: set constraints. In this step, constraint values for each of the target properties need to be given, as shown in the dialogue box of Fig. F.5.

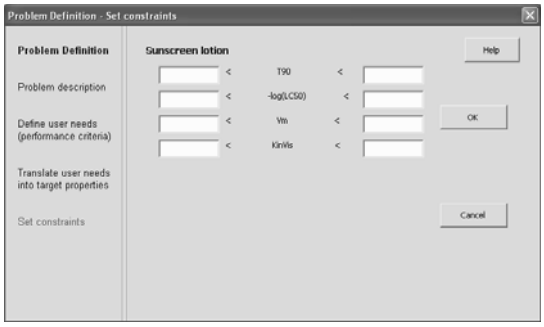


Figure F.5. The dialogue box where the user sets the constraints on the target properties.

The Help in the dialogue box of Fig. F.5 gives the units of measure for the target properties. The user needs to input at least one bound for each target property. If the constraint is of the type $\zeta < 10.0$, the lower bound is set to a very large negative value (by default). If the constraint is of the type $\zeta > 10.0$, the upper bound is set to a very large positive value (by default). For exact match of properties (for example, property $\zeta = 10.0$) the user needs to give the same value for the upper and lower bounds, and a slack of 0.05% will be applied to find the solutions as close as possible to the desired value.

F.2 Task 2: AI/AIs selection

This task corresponds to tasks S1-D2 and S1-V3 of the methodology for the design and verification of formulated products (see Chapter 5, Fig. 5.1).

Step 2.1: identify the product activity/activities. In this step the user has to select the activity/activities the product should provide. A list of the activities available in the knowledge base for each solved case study can be used (Fig. F.6). Alternatively, the user can introduce his/her own desired AI selecting the 'user defined' option at the bottom of the dialogue box of Fig. F.6.

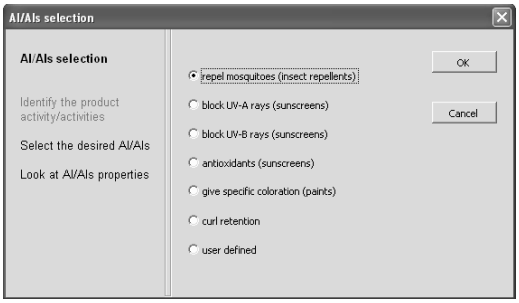


Figure F.6. The main menu of the Active Ingredient selection dialogue box.

Step 2.2: select the desired AI/AIs. In this task the AI/AIs are selected. The available AIs (in the databases) for the selected activity are displayed, as shown in Fig. F.7 for the mosquito repellent database.

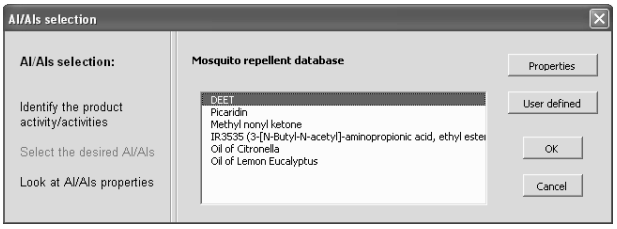


Figure F.7. The database for the Active Ingredients employed in insect repellent lotions.

Step 2.3: look at AI/AIs properties. In this step the AIs properties are checked, changed and/or missing values are added. All the new information can be saved in the selected AIs database. In Fig. F.8 the case in which DEET (from the insect repellent database) has been chosen, is shown.

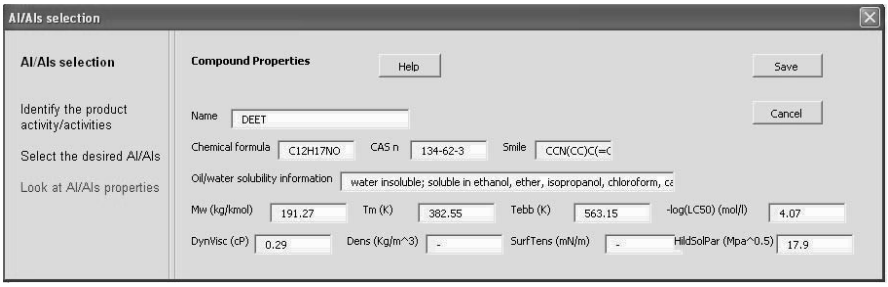


Figure F.8. The property dialogue box for one of the compounds (DEET) in the insect repellent database.

Alternatively, it is possible to introduce a new ingredient by choosing the option 'user defined' as shown in the dialogue box of Fig. F.7, and the user has the opportunity to introduce AI properties that are saved in the corresponding AIs database (if the insect repellent database was selected in the dialogue box of Fig. F.6, the new compound will be saved in this database). Another way of introducing a new ingredient to the databases is to open the Excel worksheet that contains the database and type the data/information for the new compound.
The temperature dependent data stored in the AI databases are at a temperature of 300 K (room temperature), that is the design temperature for the case studies developed in this work.

F.3 Task 3: solvent mixture design

This task corresponds to tasks S1-D3 and S1-V4 of the methodology for the design and verification of formulated products (see Chapter 5, Fig. 5.1).

Step 3.1: select solvents database. This step involves the selection/creation of a solvent database. The virtual PPD-lab hosts a solvent database library. When the user selects the 'Solvent mixture design' task, the available databases are displayed as highlighted in Fig. F.9.
The databases retrieve the solvents according to the type of solvents (alcohols, esters,…), the water solubility (water soluble, water insoluble) and the application in product design (for hair sprays, paints,…). Selecting a specific database, the user is able to access the database (an Excel worksheet) and check the data/information displayed. The user is also able to modify the information and/or add new data and compounds, and save the modifications that will be stored in the database and considered during the calculations to be performed.

Figure F.9. The dialogue box where the selection of the solvent databases is performed.

Several databases can be chosen at the same time, for a maximum total of 300 solvents. The inclusion of water in the calculations is a choice left to the user.
Users can create their own databases by selecting the 'user defined' option (Fig. F.9), that redirects the user to an empty Excel sheet where the user simply needs to insert the new information/data.
Note that in the virtual PPD-lab all the ingredient properties and the calculations based on them are at a fixed temperature of 300 K. Also solubility calculations are performed at 300 K.

Step 3.2: modelling choices. In this step the information about the models employed in the MIXD program are displayed (as shown in Fig. F.10).

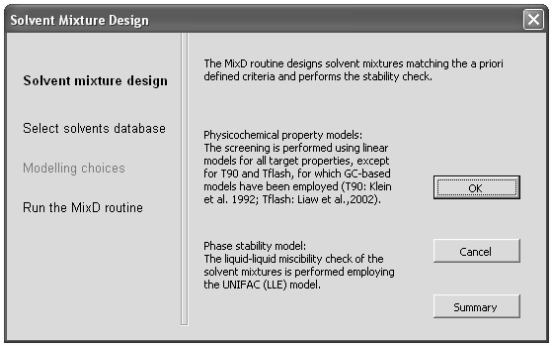


Figure F.10. The dialogue box where the information about the models employed in the MIXD program are displayed.

By clicking on 'Summary' (see the dialogue box of Fig. F.10) the user is redirected to a worksheet where data on the choices made for the performance criteria, target properties, constraints, AIs and solvent databases have been collected. Here the user is able to check the design choices and decide if to proceed with the design or modify the previous selections and/or data introduced.

Step 3.3: run MIXD. In this step it is possible to launch the mixture design program (MIXD), as shown in Fig. F.11.

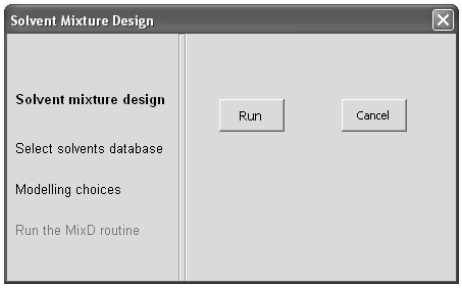


Figure F.11. The dialogue box from that the user can launch the mixture design program.

When the calculations are over, the user is able to see the results by clicking on 'See results' in the dialogue box of Fig. F.12. The generated output file contains, at first, the identity of the mixtures matching the constraints defined in task 1: compounds forming the mixture, composition and cost. Then, the mixtures properties are listed for each mixture. Finally, the mixture stability is checked. If the evaporation time T_{90} or the flash temperature T_f (properties described by non-linear models, see dialogue box of Fig. F.10) were selected in the problem design task, information about these two properties are given before the stability test. That is, for each mixture, the T_{90} value is reported, and a string stating if the value matches the *a priori* defined target on T_{90} is also displayed. The same is done for the flash temperature.

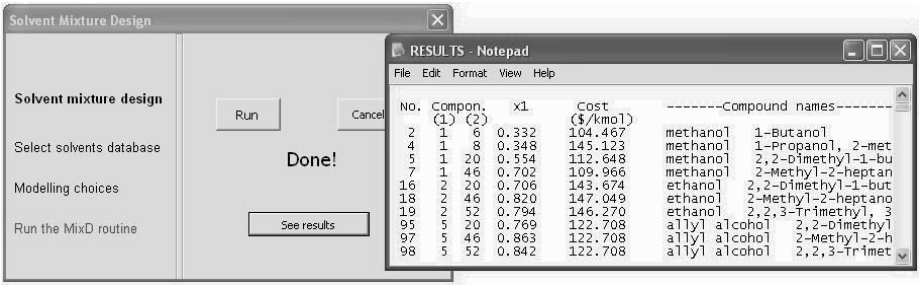


Figure F.12. The results from the mixture design program (MIXD) are displayed.

F.4 Additives selection task

This task corresponds to tasks S1-D4 and S1-V5 of the methodology for the design and verification of formulated products.

Step 4.1: qualities to enhance. In this step the product qualities to enhance/add are selected. Many additives can be added to a formulated product. The most common additives for consumer oriented products are the aroma compounds.
The 'Additive selection' task hosts the aroma database, to improve the scent of the designed formulated product, and several user-defined additives databases, where the user can define his/her own additive (Fig. F.13).

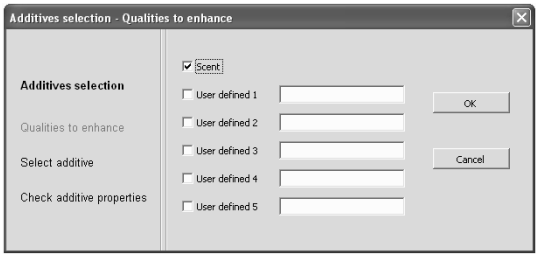


Figure F.13. The main menu of the additives selection dialogue box.

If the user selected 'User defined 1' from the dialogue box of the quality to enhance (see Fig. F.14), it will be possible to input another product quality to enhance, for example, spread-ability (quality to enhance for a paint formulation, for instance).

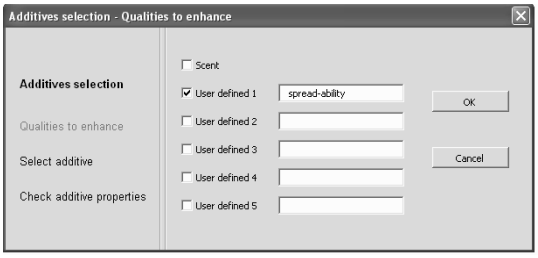


Figure F.14. The choice of a user defined quality to enhance from the main menu of the additive selection task.

Step 4.2: select additive. In this step the user can search the selected database or create a new database.
The aroma database (Fig. F.15) contains 350 compounds that are employed in the cosmetic and food industry. The user can enter the database by clicking on 'View database', and/or search the database (see dialogue box of Fig. F.15). Aroma can be searched in terms of name, CAS number, smile, smell class and solubility information. Aromas are grouped in terms of smell classes (green, fruit, etc), but each aroma copound has its own particular scent (for the fruit class, the scent could be:

almond, peach, strawberry, and so on). The search of aroma compound is performed according to the general smell class, but information about the particular scent of a compound is reported in the 'Properties' dialogue box (step 4.3 and Fig. F.17).
For the example shown in Fig. F.15, the user searched for aroma compounds in the class of 'aromatic leaf and seeds', and the search returned all the aroma compounds belonging to this class.

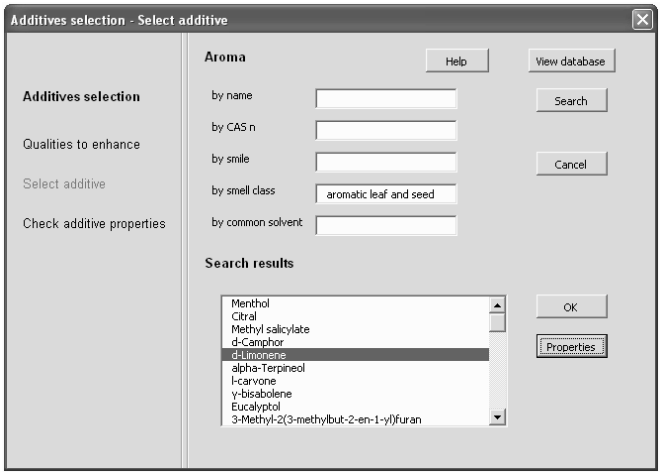


Figure F.15. The search in the aroma database for aroma compounds belonging to the 'aromatic leaf and seeds' class.

The Help in the dialogue box of Fig. F.15 informs the user about all the possible smell classes, and also about all the possible smells.
In the case the user decided to introduce a user defined quality, as in the dialogue box of Fig. F.14, it will be possible to choose the number of additives to add in order to enhance the selected product quality. In Fig. F.16 an example is shown: only one additive (sodium dioctyl sulfosuccinate) is introduced to enhance the spread-ability in the case of a paint formulation.

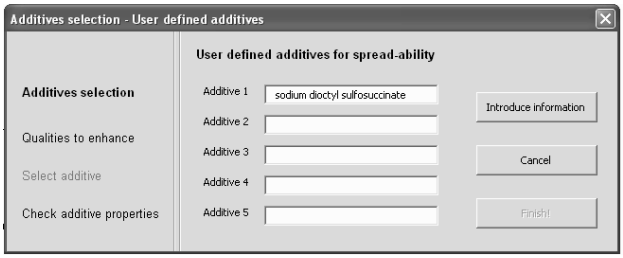


Figure F.16. The introduction of one user defined additive that enhances the spread-ability.

Step 4.3: check additive properties. The properties of the additives are checked and/or introduced in this step.

If the aroma database was searched, and one of the compounds in the list of aroma compounds matching the target has been selected, the command button 'Properties' (in Fig. F.15) will be enabled and the user will be able to check the aroma compound properties (as in Fig. F.17).

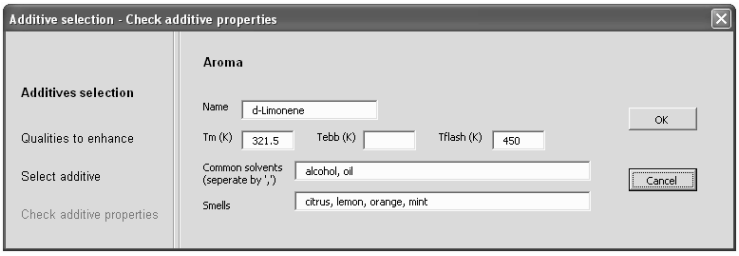


Figure F.17. The dialogue box that displays the properties of one aroma compound.

In case a user defined compound was introduced as shown in the dialogue box of Fig. 7.17, the 'Introduce information' command button will be enabled and the user can introduce the data/information for that compound (Fig. F.19).

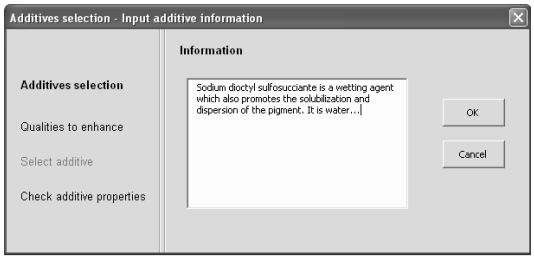


Figure F.18. Introduction of the information for the user defined compound nº 1 for spread-ability (sodium dioctyl sulfosuccinate).

The design/verification can be terminated clicking the 'Ok' button in Fig F.15 (after the aroma compound has been selected) or clicking the 'Finish' button in Fig. F.16.

The virtual PPD-lab will redirect the user to the summary worksheet, where he/she can check his/her design/verification results.

G

Hair spray case study: intermediate results

The intermediate results for the hair spray case study are given in this appendix. They consist of:

- Binary phase diagrams for the polymer-solvent systems that give phase splits around the temperature of interest (300 K), (G.1);
- Ternary phase diagrams for the polymer-solvent binary mixture systems that gave phase splits at the temperature of interest (300 K), (G.2).

G.1 Binary phase diagrams

The binary phase diagrams are shown in Figs. G.1-G.3:

- Fig. G.1: phase diagrams involving polymer *P*1;
- Fig. G.2: phase diagrams involving polymer *P*2;
- Fig. G.3: phase diagrams involving polymer *P*3.

For some systems the phase equilibria could not be predicted for the entire temperature range because the calculated polymer activity coefficients at these conditions (high or low temperatures) become too sensitive to the regressed parameters. As a consequence, it was not always possible to obtain the polymer-solvent system phase behaviour that resembles the known type UCST (Upper Critical Solution Temperature), LCST (Lower Critical Solution Temperature), or, hour glass and closed loop.

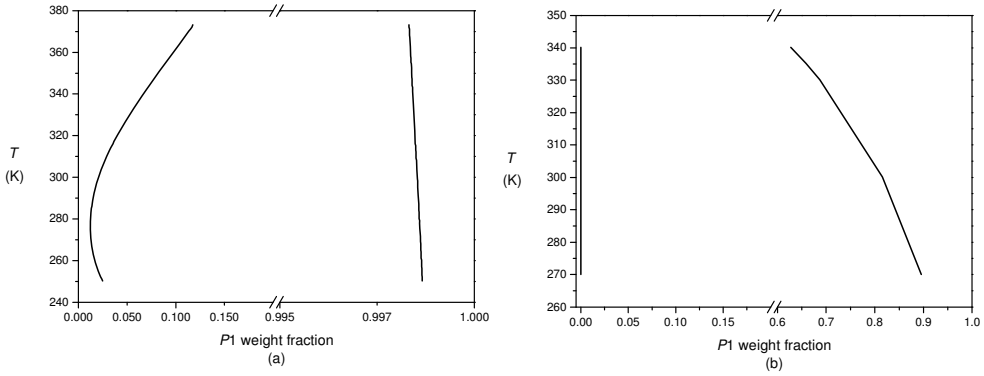


Figure G.1. Predicted phase behaviour for the systems constituted of polymer *P*1: (a) (*P*1-*A*) and (b) (*P*1-*B*).

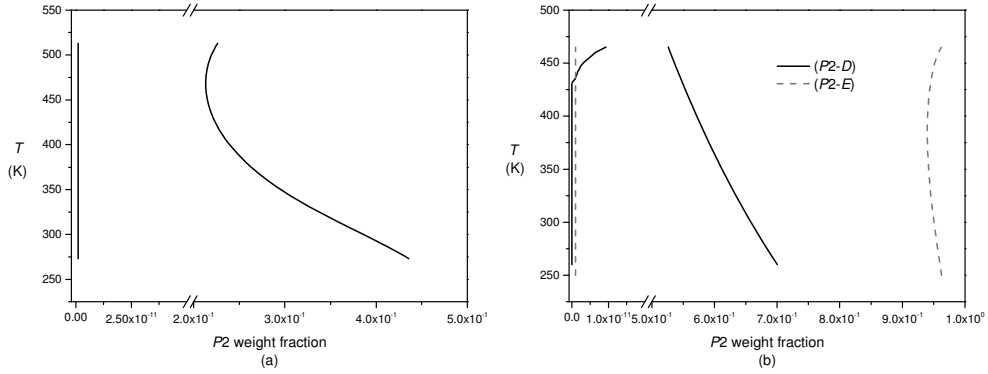


Figure G.2. Predicted phase behaviour for the systems constituted of polymer *P*2: (a) (*P*2-*A*); (b) (*P*2-*D*) and (*P*2-*E*).

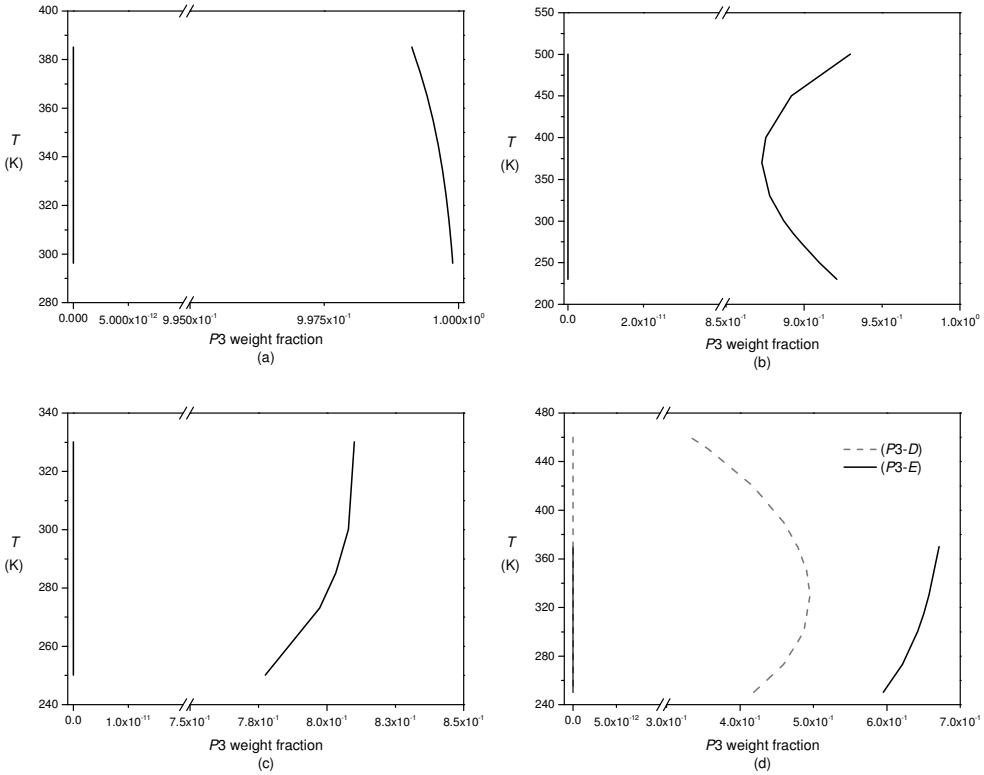


Figure G.3. Predicted phase behaviour for the systems: (a) (*P*3-*A*); (b) (*P*3-*B*); (c) (*P*3-*C*); (d) (*P*3-*D*) and (*P*3-*E*).

G.2 Ternary phase diagrams

The ternary phase diagrams are shown in Figs. G.4 –G.10:

- Figs. G.4-G.5: phase diagrams involving polymer *P*1;
- Figs. G.6-G.7: phase diagrams involving polymer *P*2;
- Figs. G.8-G.10: phase diagrams involving polymer *P*3.

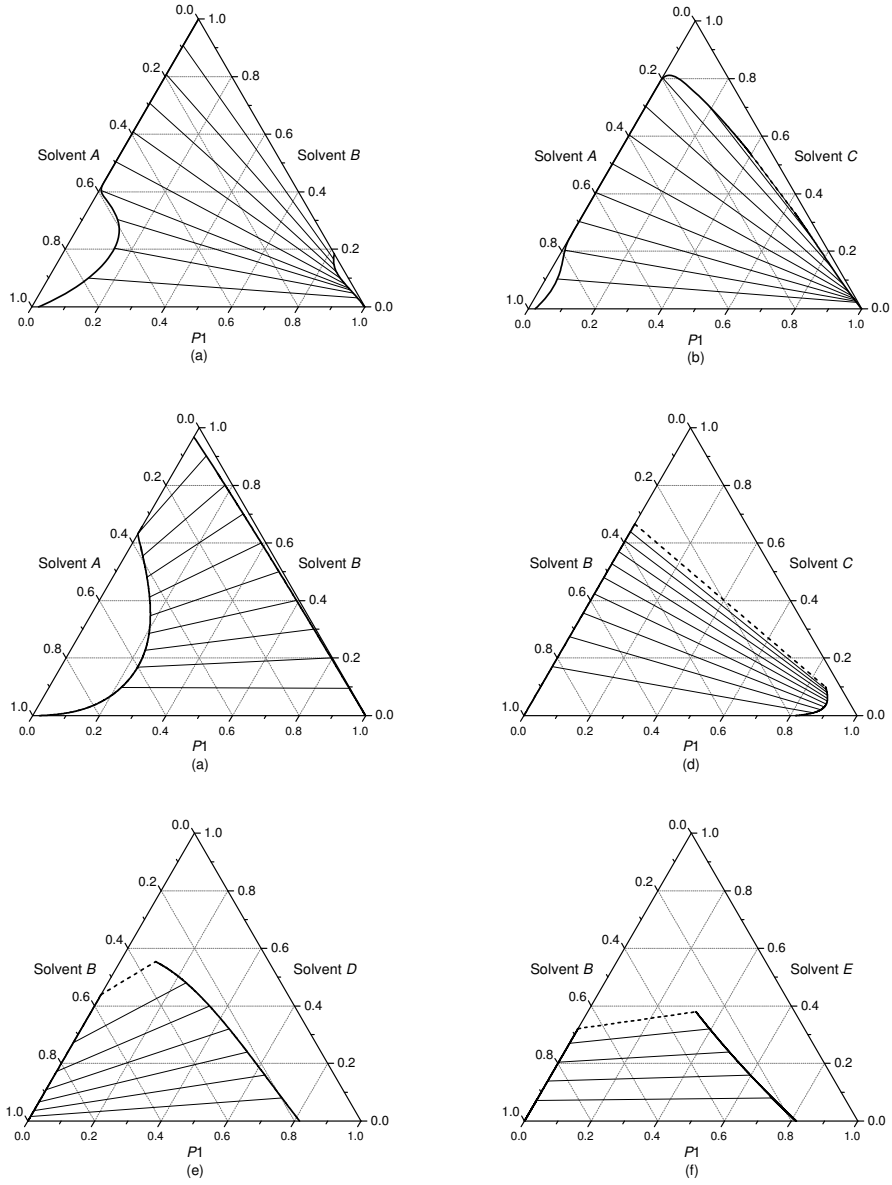


Figure G.4. Weight based ternary phase equilibrium diagrams for: (a) (*P*1-*A*-*B*); (b) (*P*1-*A*-*C*); (c) (*P*1-*A*-*D*); (d) (*P*1-*B*-*C*); (e) (*P*1-*B*-*D*); (f) (*P*1-*B*-*E*).

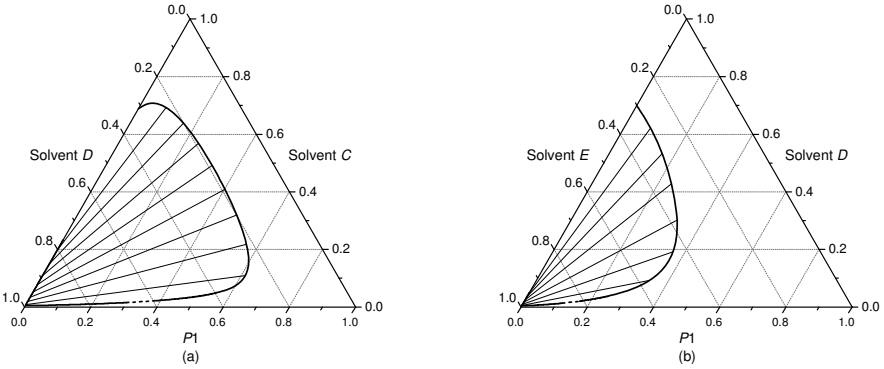


Figure G.5. Weight based ternary phase diagrams for (a) (*P*1-*C*-*D*) and (b) (*P*1-*D*-*E*).

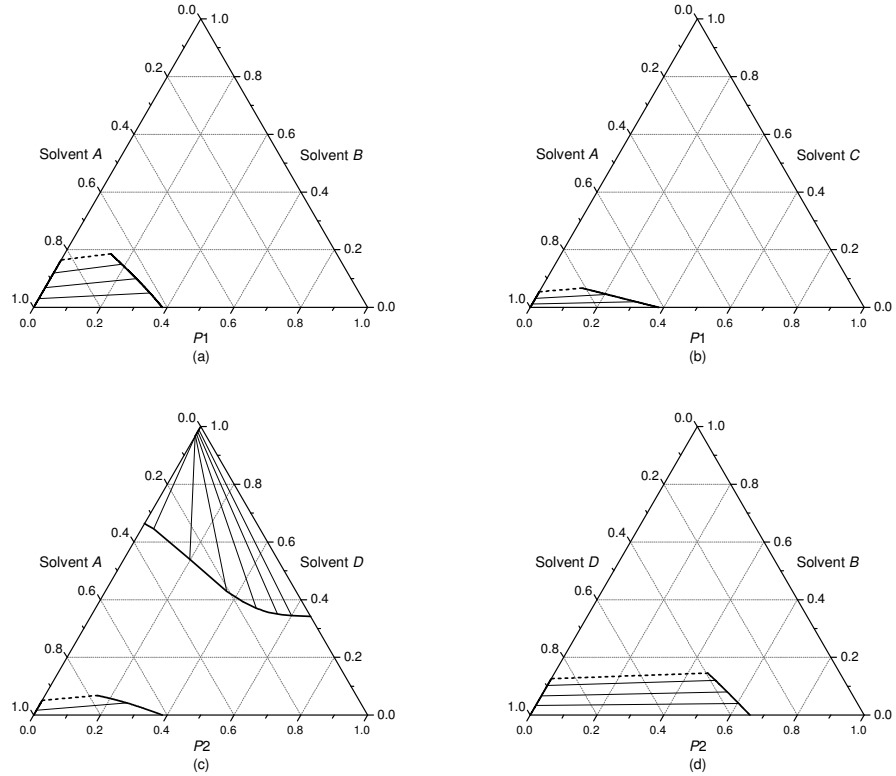


Figure G.6. Weight based ternary phase diagrams for: (a) (*P*2-*A*-*B*); (b) (*P*2-*A*-*C*); (c) (*P*2-*A*-*D*); (d) (*P*2-*B*-*D*).

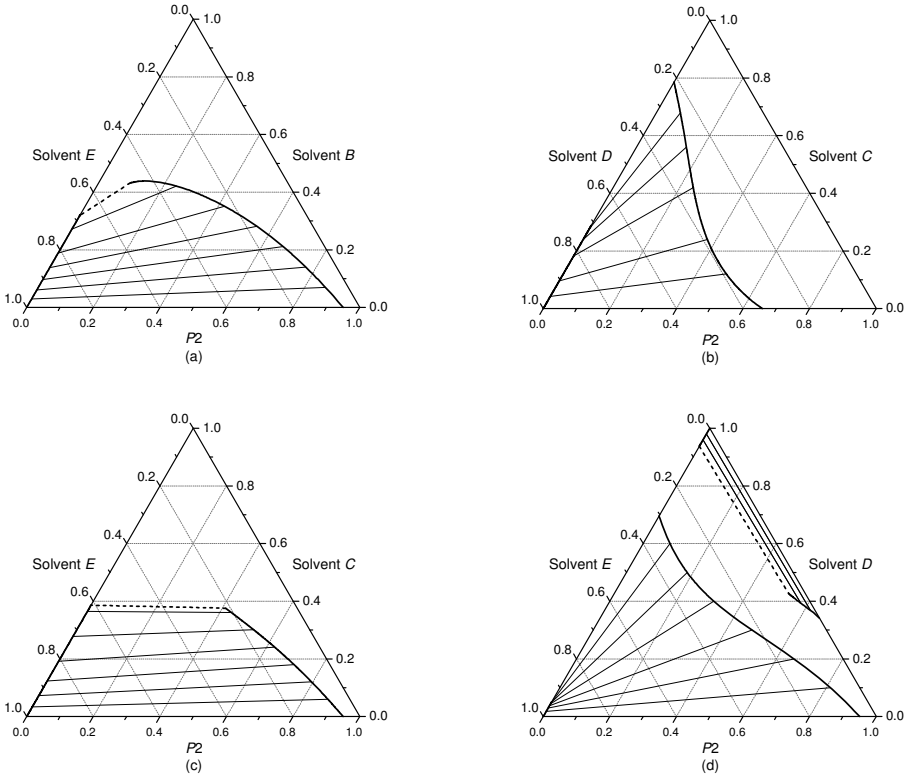


Figure G.7. Weight based ternary phase diagrams for: (a) (*P*2-*B*-*E*); (b) (*P*2-*C*-*D*); (c) (*P*2-*C*-*E*); (d) (*P*2-*D*-*E*).

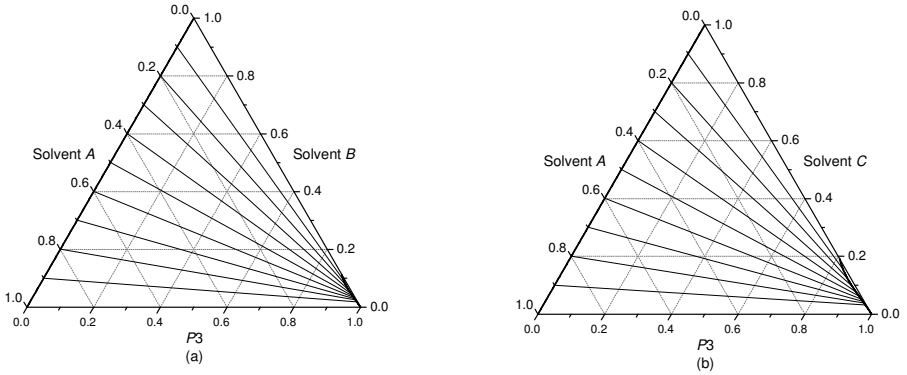


Figure G.8. Weight based ternary phase diagrams for (a) (*P*3-*A*-*B*) and (b) (*P*3-*A*-*C*).

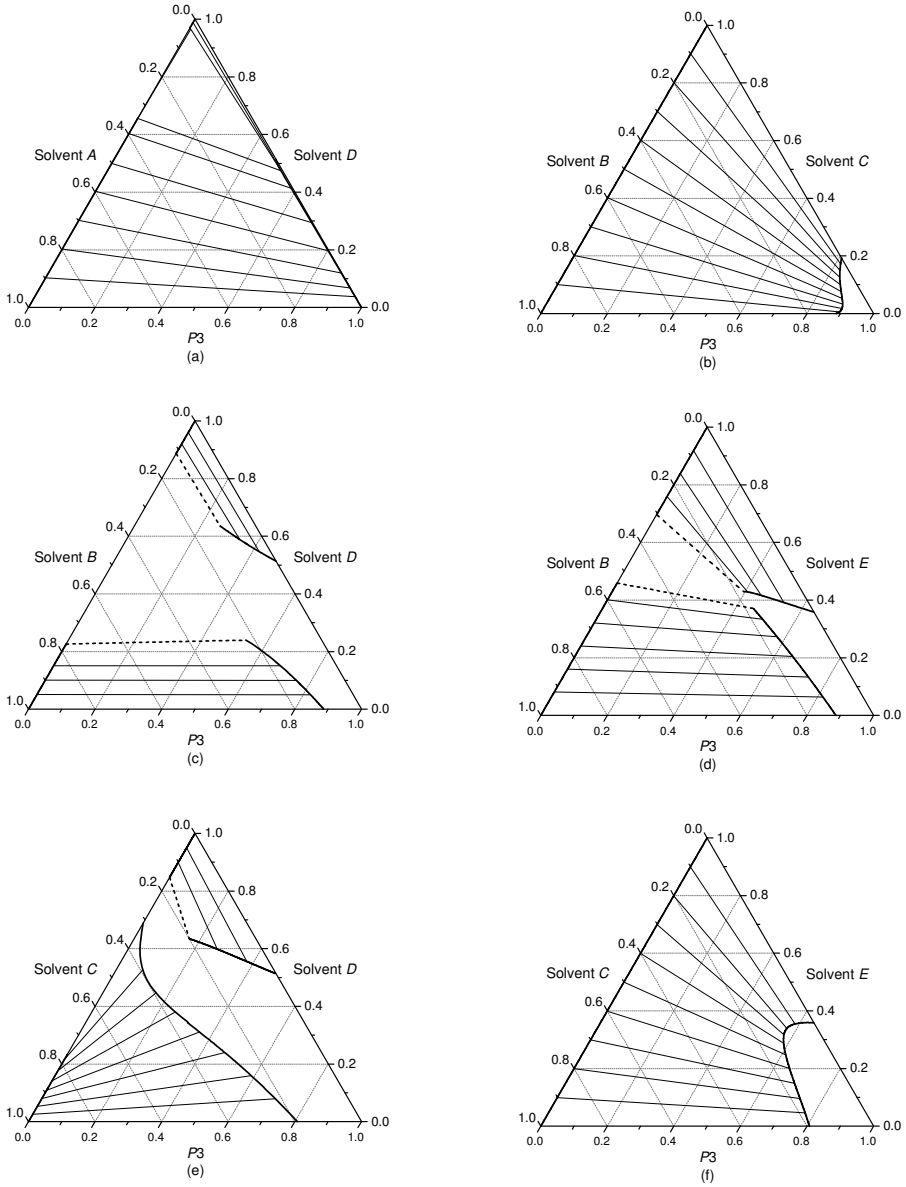


Figure G.9. Weight based ternary phase diagrams for: (a) (*P*3-*A*-*D*); (b) (*P*3-*B*-*C*); (c) (*P*3-*B*-*D*); (d) (*P*3-*B*-*E*); (e) (*P*3-*C*-*D*); (f) (*P*3-*C*-*E*).

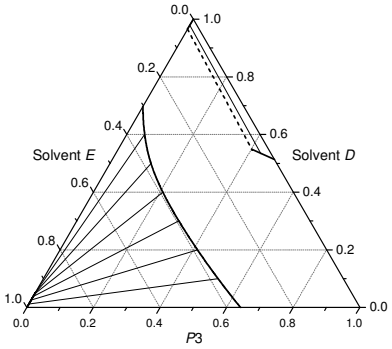


Figure G.10. Weight based ternary phase diagram for the system (*P*3-*D*-*E*).

Nomenclature

a	=	activity
at	=	slope of the straight line calculated for the geometrical solution of the *TPD* condition
ac	=	atom occurrences in the molecular structure (CI method)
a_{ij}	=	isothermal group/segmental interaction parameter between groups i and j
A_w	=	Van der Waals surface area
A	=	molar surface area (Suarez *et al.*, 1989)
$A^{O\&E}$	=	Orrick and Erbar (1974) model parameter
AC	=	atom contribution (CI method)
b_{ij}	=	temperature dependent segmental interaction parameter
$B^{O\&E}$	=	Orrick and Erbar (1974) model parameter
bt	=	intercept of the straight line calculated for the geometrical solution of the *TPD* condition
C	=	group contribution of the first-order M&G
C^{DOF}	=	temperature-dependent molecular external degrees of freedom parameter(GC-Flory EoS)
C_{T0}	=	group parameter for the calculation of C^{DOF} (GC-Flory EoS)
C_T	=	group parameter for the calculation of C^{DOF} (GC-Flory EoS)
C^0	=	group parameter for the calculation of C^{DOF} (GC-Flory EoS)
D	=	group contribution of the second-order M&G
Dm	=	dipolar moment
E	=	group contribution of the third-order M&G
$F()$	=	function of ()
ΔG^{mix}	=	delta Gibbs energy of mixing
G^E	=	excess Gibbs energy
h	=	binary variable for the M&G method for property prediction
LC_{50}	=	lethal concentration that causes the death of the 50% of a fathead minnow population
M	=	group occurrence of the second-order
Mw	=	molecular weight
Mwd	=	molecular weight distribution
nD	=	refractive index
n_H	=	number of hydrogen atoms attached to a carbon atom
np	=	proportional constant (Cao *et al.*, 1993)
N	=	group occurrence of the first-order
N^{tot}	=	total number of regressed data points
$N_{S\&R}$	=	Sastri and Rao (1992) model parameters
NB	=	total number of bonds in a molecule
NK	=	total number o missing fragments k in a molecule
NV	=	total number of vertexes in a molecule
O	=	group occurrence of the third-order
$objF$	=	objective function
P_C	=	critical pressure
P^{sat}	=	vapour pressure
q	=	surface area parameter
Q	=	group area parameter, calculated from the Van der Waals surface area A_w
r	=	volume parameter
R	=	universal constant of gases
R	=	group volume parameter, calculated from the Van der Waals volume V_w
R^2	=	square residual
s	=	adjustable parameter for the Connectivity Index based method (CI method)
t	=	adjustable parameter for the Connectivity Index based method (CI method)
T	=	temperature
T_b	=	boiling temperature (at atmospheric pressure)

T_C = critical temperature
T_f = closed-cup flash point
T_m = melting temperature
T_{ref} = reference temperature
T_{10} = time at which the 10% by weight of a pure compound or mixture has evaporated
T_{50} = time at which the 50% by weight of a pure compound or mixture has evaporated
T_{90} = time at which the 90% by weight of a pure compound or mixture has evaporated
u = constant for the Connectivity Index based method (CI method)
ΔU_{ij} = energy of interaction between groups i and j
V = molar volume (of the liquid bulk)
V_c = critical volume
V^{FV} = free volume (molar volume-hard core volume)
$\tilde{V}$ = reduced volume (ratio between v_i and v^*)
V^* = hard core volume for
V^{**} = normalized hard core volume (v^*/1.448)
V_w = Van der Waals volume
w = weight fraction
x = molar fraction
$x_{i,s}$ = surface molar fraction
$x_{i,b}$ = bulk molar fraction
X = molar fraction of segments (FV-UNIQUAC)
y = binary variable for the M&G method for property prediction
z = coordination number (number of nearest neighbours molecules)

Greek symbols

β^k = bond index for bond k (CI method)
γ = molar base activity coefficient
$\gamma_{i,s}$ = surface molar base activity coefficient
$\gamma_{i,b}$ = bulk molar base activity coefficient
Γ = group/segment activity coefficient
$\Gamma_{i,j}$ = group/segment j activity coefficient in the mixture of segments of the molecule i
δ^v = atomic index (CI method)
$\Delta\varepsilon_{ji}$ = GC-Flory EoS energetic parameter
$\Delta\eta_B$ = Sastri and Rao (1992) group contribution parameter
$\Delta\eta_{cor}$ = Sastri and Rao (1992) group contribution parameter
ΔN^{cor} = Sastri and Rao (1992) group contribution parameter
$\Delta N^{S\&R}$ = Sastri and Rao (1992) group contribution parameter
ε = dielectric constant of compound i
ε^0_{ji} = interaction energy
ε_{ji} = GC-Flory EoS energetic parameter
ζ = generic property
$\underline{\zeta}$ = vector of generic (target) properties
η = viscosity
η_B = Sastri and Rao (1992) model parameter
η_{cor} = Sastri and Rao (1992) model parameter
θ = surface area fraction
θ_{ij} = interaction and surface area parameter (Cao *et al.*, 1993)
Θ_i = segment surface fraction
ξ = Staverman-Guggenheim correction
ρ = liquid density
σ = surface tension
τ_{ij} = exponential form of the interaction parameter (a_{ij})
$\upsilon_{k,i}$ = number of groups/fragments/segments of type k in the molecule i
φ = volume fraction
φ^{FV} = free volume fraction

${}^{v}\chi^{0}$	=	zero-order connectivity index (CI method)
${}^{v}\chi^{1}$	=	first-order connectivity index (CI method)
$\underline{\Psi}$		vector of performance criteria
Ω	=	weight base activity coefficient
ω	=	acentric factor

Acronyms

AI	=	Active Ingredient
AAD	=	Average Absolute Error
ER	=	Evaporation Rate
HB	=	Hydrogen Bond
LLE	=	Liquid-Liquid Equilibrium
MED	=	Minimal Erythematic Dose
NA	=	Number of Atoms
NB	=	Number of total Bonds
Nbi	=	Number of internal bonds
Nbe	=	Number of external bonds
NC	=	Number of Compounds
NE	=	Number of Experiments
NG	=	Number of total Groups
*NG*1	=	Number of total Groups
*NG*2	=	Number of total Groups
*NG*3	=	Number of total Groups
NK	=	Number of missing fragments in a molecule/groups in a molecule
NP	=	Number of Properties
NS	=	Number of Segments
NV	=	Number of Vertices
PDI	=	Polidispersity Index
PI	=	Performance Index
PS	=	Protected skin
RD	=	Relative Deviation
SD	=	Standard Deviation
SPF	=	Sun Protection Factor
TPD	=	Tangent Plane Distance
US	=	Unprotected Skin
VLE	=	Vapour-Liquid Equilibrium
VOCs	=	Volatile Organic Compounds

References

Abildskov J. and Kontogeorgis G. M. (2004). Chemical product design. A new challenge of applied thermodynamics. *Chemical Engineering Research and Design*. 82(A11), 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.

Al-Hayan M. N. M. and Abdul-latif A.-H. M. (2006). Excess molar volumes and viscosities of (1,1,2,2-tetrabromoethane+1-alkanols) at T = (293.15 and 303.15) K. *Journal of Chemical Thermodynamics.* 38, 68-74.

Ali A., Khan S., Hyder S. and Tariq M. (2007). Interactions of some a-amino acids with tetra-n-alkylammonium bromides in aqueous medium at different temperatures. *Journal of Chemical Thermodynamics.* 39, 613-620.

Al-Kandary J. A., Al-Jimaz A. S. and Abdul-Latif A.-H. M. (2006). Densities, viscosities, and refractive indices of binary mixtures of anisole with benzene, methylbenzene, ethylbenzene, propylbenzene, and butylbenzene at (293.15 and 303.15) K. *Journal of Chemical Engineering Data*. 51, 99-103.

American Society for Testing and Materials (ASTM). (1987). Evaporation rates of volatile liquids by shell thin-film evaporometer. Method #D 3539-87.

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.

Badolo A., Ilboudo-Sanogo E., Patoin Ouédraogo A. and Costantini C. (2004). Evaluation of the sensitivity of Aedes aegypti and Anopheles gambiae complex mosquitoes to two insect repellents: DEET and KBR 3023. *American Journal of Tropical Medicine and Hygiene*. 9(3), 330-334.

Bagajewicz M. J. (2007). On the role of microeconomics, planning, and finances in product design. *AIChE Journal*. 53(12), 3155-3170.

Baker L. E., Pierce A. C. and Luks K. D. (1982). Gibbs energy analysis of phase equilibria. *Journal of Petroleum Science and Engineering, AIME.* 22(5), 731-738.

Bird R. B., Stewart W. E. and Lightfoot E. N. (2002). *Transport phenomena*. Wiley, New York, NY.

Bleazard J. G., Sun T. F., Johnson R. D., DiGiulio R. M. and Teja A. S. (1996). The transport properties of seven alkanediols. *Fluid Phase Equilibria*. 117, 386-393.

Bogdanic G. and Fredenslund, A. (1994). Revision of the Group-Contribution Flory Equation of State for phase equilibria calculations in mixtures with polymers. *Industrial and Engineering Chemistry Research*. 33(5), 1331-1340.

Bogdanic G. and Vidal J. (2000). A segmental interaction model for liquid–liquid equilibrium calculations for polymer solutions. *Fluid Phase Equilibria*. 173, 241-252.

Bogdanic G. (2001). The FV-UNIQUAC segmental interaction model for liquid–liquid equilibrium calculations for polymer solutions. Part 2. Extension to solutions containing polystyrene. *Fluid Phase Equilibria*. 191, 49-57.

Bondi A. (1968). *Physical Properties of molecular crystals, liquid and glasses*. Wiley, New York, NY.

Browarzik D. (2005). Phase-equilibrium calculations for non-aqueous and aqueous associating systems using continuous thermodynamics. *Fluid Phase Equilibria*. 230, 143-152.

Browarzik D. (2007). Continuous thermodynamics of binary associating systems. *Fluid Phase Equilibria.* 254, 174-187.

Camarda K. V. and Maranas C. D. (1999). Optimization in polymer design using connectivity indices. *Industrial and Engineering Chemistry Research*. 38, 1884-1892.

Cao W., Knudsen K., Fredenslund Aa. and Rasmussen P. (1993). Simultaneous correlation of viscosity and vapour-liquid equilibrium data. *Industrial Engineering Chemistry research.* 32, 2077-2087.

Cháfer A., Lladosa E., de la Torre J. and Burguet M. C. (2008). Study of liquid–liquid equilibrium of the systems isobutyl acetate + acetic acid + water and isobutyl alcohol + acetic acid + water at different temperatures. *Fluid Phase Equilibria*. 271, 76-81.

Charpentier J. C. (2002). The triplet "molecular processes–product–process" engineering: the future of chemical engineering? *Chemical Engineering Science*. 57, 4667-4690.

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.

Charpentier J. C. (2010). Among the trends for a modern chemical engineering, the third paradigm: the time and length multiscale approach as an efficient tool for process intensification and product design and engineering. *Chemical Engineering Research and Design.* 88, 248-254.

Chemmangattuvalappil N. G., Solvason C. C., Bommareddy S. and Eden M. R. (2010). Combined property clustering and GC^+ techniques for process and product design. *Computers and Chemical Engineering*. 34(5), 582-591.

Chen F., Fredenslund Aa. and Rasmussen P. (1990). Group-Contribution Flory Equation of State for vapour-liquid equilibria in mixtures with polymers. *Industrial and Engineering Chemistry Research*. 29, 875-882.

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.

Churi N. and Achenie L. E. K. (1996). Novel mathematical programming model for computer-aided molecular design. *Industrial and Engineering Chemistry Research*. 35, 3727-3739.

Committee on Chemical Engineering Frontiers. (1988). *Frontiers in chemical engineering: research needs and opportunities*. National Academy Press, Washington, WA.

Constantinou L. and Gani R. (1994). A new group-contribution method for the estimation of properties of pure compounds. *AIChE Journal.* 40(10), 1697-1709.

Constantinou L., Bagherpour K, Gani R., Klein J. A. and Wu D. T. (1996). Computer aided product design: problem formulations, methodology and applications. *Computers and Chemical Engineering*. 20(6/7), 685-702.

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 (20), 7940-7954.

Conte E., Morales-Rodriguez R. and Gani R. (2009a). The virtual product-process design laboratory as a tool for product development. In *Proceedings of 19th European Symposium on Computer-Aided Process Engineering*. 29, 249-254, Cracow, Poland, June 2009.

Conte E., Morales-Rodriguez R. and Gani R. (2009b). The virtual product-process design laboratory as a tool for design and analysis of formulations. In *Proceedings of 10th International Symposium on Process System Engineering*. 825-830, Salvador de Bahia, Brazil, August 2009.

Conte E., Gani R. and Ng K. M. (2010a). Design of formulated products: A systematic methodology. Submitted to *AIChE Journal*.

Conte E., Gani R. and Malik T. I. (2010b). The virtual product-process design laboratory to manage the complexity in formulation design. Submitted to *Fluid Phase Equilibria*.

Costa R., Moggridge G. D. and Saraiva P. M. (2006). Chemical product engineering: an emerging paradigm within chemical engineering. *AIChE Journal*. 52(6), 1976-1986.

Coutinho J. A. P., Vilela T., Pereira P., Pessoa P., Santos M. M. M. and Kontegeorgis G. M. (2005). Process engineering versus product engineering. A case study on volatile organic compounds removal. *Chemical Engineering Research and Design*. 83(A4), 352-356.

Cruz R. da C., Martins R. J., De M. Cardoso M. J. E, and Barcia O. E. (2002). Measurement of the viscosity of binary and ternary systems containing benzaldehyde, toluene, and benzyl alcohol at 293.15 K and 0.1 MPa. *Journal of Chemical Engineering Data*. 47, 927-931.

Cussler E. L. and Moggridge G. D. (2001). *Chemical Product Design*. Cambridge University Press, Cambridge, ME.

Cussler E. L. and Wei J. (2003). Chemical Product Engineering. *AIChE Journal*. 49(5), 1072-1075.

Ćwiklińska A., Dzikowski T., Szychowski D., Kinart W. J. and Kinart C. M. (2007). Viscosities of binary mixtures of some n-ethoxyethanols with ethyl tert-butyl ether at T = (293.15, 298.15, and 303.15) K. *Journal of Chemical Thermodynamics.* 39, 1325-1330.

Danckwerts P. V. (1996). Science in chemical engineering. *Chemical Engineering*. 200, CE155-CE159.

Danner R. P. and High M. S. (1992). *Polymer Solution Handbook, AIChE*. Pennsylvania State University, PA.

Das M. and Roy M. N. (2006). Studies on thermodynamic and transport properties of binary mixtures of acetonitrile with some cyclic ethers at different temperatures by volumetric, viscometric, and interferometric techniques. *Journal of Chemical Engineering Data*. 51, 2225-2232.

Daubert T. E. and Danner R. P. (1985). *Data compilation tables of properties of pure compounds*. DIPPR/AIChE.

Dean J. A. (1992). *Lange's Handbook of Chemistry*. McGraw-Hill, New York, NY.

Debboun M., Frances S. P. and Strickman D. (2006). *Insect repellents: principles, methods, and use.* CRC Press, Boca Raton, FL.

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(5), 483-496.

Derringer G. C. and Markham R. L. (1985). A computer-based methodology for matching polymer structure with required properties. *Journal of Applied Polymer Science*. 30, 4609-4617.

Duvedi A. P. and Achenie L. E. K. (1996). Designing environmentally safe refrigerants using mathematical programming. *Computers Engineering Science*. 51, 3727-3739.

Eden M. R. (2003). *Property based process and product synthesis and design*. Ph.D. Dissertation, Technical University of Denmark, Denmark.

Eden M. R., Jørgensen S. B., Gani R. and El-Halwagi M. M. (2004). A novel framework for simultaneous separation process and product design. *Chemical Engineering and Processing*. 43, 595-608.

Elbro H. S., Fredenslund Aa. and Rasmussen P. (1990). A new simple equation for the prediction of solvent activities in polymer solutions. *Macromolecules*. 23, 4707-4714.

Elbro H. S., Fredenslund Aa. 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(12), 2576-2582.

El-Halwagi M. M., Glasgow I. M., Qin X. and Eden M. (2000). Property integration: componentless design techniques and visualization tools. *AIChE Journal*. 50(8), 1854-1869.

Eljack F. T., Eden M. R., Kazantzi V., Qin X. and El-Halwagi M. M. (2007). Simultaneous process and molecular design - Property based approach. *AIChE Journal*. 53(5), 1232-1239.

Favre E., Marchal-Heusler L. and Kind M. (2002). Chemical product engineering: research and educational challenges. *Transactions of IChemE*. 80A, 65-74.

Food and Drug Administration. (1978). Sunscreen drug products for over-the counter human drugs. Proposed safety, effective and labeling conditions. *Federal Register*. 43, 38259-38269.

Frances S. P., Marlow R. M., Jansen C. C., Huggins R. L. and Cooper R. D. (2005). Laboratory and field evaluation of commercial repellent formulations against mosquitoes (Diptera: Culicidae) in Queensland, Australia. *Australian Journal of Entomology*. 44, 431-436.

Fredenslund Aa., Jones R. L. and Prausnitz A. M. (1975). Group-contribution estimation of activity coefficient in non ideal mixtures. *AIChE Journal*. 21(6), 1086-1099.

Friedler F., Fan L. T., Kalotai L. and Dallos A. (1998). A combinatorial approach for generating candidate molecules with desired properties base don group contribution. *Computers and Chemical Engineering*. 22(6), 809-817.

Fung K. Y. and Ng K. M. (2003). Product-centered processing: pharmaceutical tablets and capsules. *AIChE Journal*. 49(5), 1193-1215.

Gani R., Nielsen B. and Fredenslund Aa. (1991). A group contribution approach to computer-aided molecular design. *AIChE Journal*. 37(9), 1318-1332.

Gani R. and Brignole E. S. (1993). Molecular design of solvents for liquid extraction based on UNIFAC. *Fluid Phase Equilibria*. 13, 331-340.

Gani R. and Pistikopoulos E. N. (2002). Property modeling and simulation for product and process design. *Fluid Phase Equilibria*. 194-197, 43-59.

Gani R. (2004a). Chemical product design: challenges and opportunities. *Computers and Chemical Engineering*. 28(12), 2441-2457.

Gani R. (2004b). Computer-aided methods and tools for chemical product design. *Chemical Engineering Research and Design.* 82(A11), 1494-1504.

Gani R., Harper P. M. and Hostrup M. (2005a). Automatic creation of missing groups through connectivity index for pure-component property prediction. *Industrial and Engineering Chemistry Research.* 44(18), 7262-7269.

Gani R., Jimenez-Gonzalez C. and Constable D. J. C. (2005b). Method for selection of solvents for promotion of organic reactions. *Computers and Chemical Engineering*. 29, 1661-1676.

Gascón I., Mainar A. M., Royo F. M. and Urieta J. S. (2000). Experimental viscosities and viscosity predictions of the ternary mixture (cyclohexane+1,3-dioxolane+2-butanol) at 298.15 and 313.15 K. *Journal of Chemical Engineering Data*. 45, 751-755.

Giner B., Artigas H., Haro M., Lafuente C. and López M. C. (2006). Viscosities of binary mixtures of 1,3-dioxolane or 1,4-dioxane with isomeric chlorobutanes. *Journal of Molecular Liquids*. 129, 176-180.

Gramajo de Doz M. B., Cases A. M. and Sólimo H. N. (2006). (Liquid + liquid) equilibria for the quaternary system water + methyl *tert*-butyl ether + ethanol + benzene at 303.15K. *Fluid Phase Equilibria*. 249, 109-114.

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. (2007). *Hansen Solubility Parameters: A User's Handbook.* CRC Press, Boca Raton, FL.

Hao W., Elbro H. S. and Alessi P. (1992). *Polymer solution data collection part 2, 3. DECHEMA Chemistry Data Series 14, 2-3.* DECHEMA, Frankfurt, Germany.

Harper P. M., Gani R., Kolar P. and Ishikawa T. (1999). Computer-aided molecular design with combined molecular modelling and group contribution. *Fluid Phase Equilibria*. 158-160, 337-347.

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.

Hill M. (2004). Product and Process Design for Structured Products. *AIChE Journal*. 8(50), 1656-1661.

Hill M. (2009). Chemical product engineering-The third paradigm. *Computers and Chemical Engineering*. 33, 947-953.

Holten-Andersen J., Rasmussen P. and Fredenslund Aa. (1987). Phase equilibria of polymer solutions by group contribution. 1. Vapour-liquid equilibria. *Industrial and Engineering Chemistry Research*. 26, 1382-1390.

Horvath A. L. (1992). *Molecular Design*. Elsevier, Amsterdam, The Netherlands.

Hwang I.-C., Park S.-J., Kim K.-L. and Han K.-J. (2007a). Liquid-Liquid Equilibria for Binary System of Ethanol + Hexadecane at Elevated Temperature and the Ternary Systems of Ethanol + Heterocyclic Nitrogen Compounds + Hexadecane at 298.15 K. *Journal of the American Chemical Society*. 52, 1019-1924.

Hwang I.-C., Han J.-H. and Park S.-J. (2007b). Binary LLE for Propyl Vinyl Ether (PVE) + Water, Ternary LLE for PVE + Methanol or Ethanol + Water at 298.15 K, and V^E and ΔR at 293.15 K for the Mixture of PVE + Ethanol + 2,2,4-Trimethylpentane. *Journal of the American Chemical Society*. 52, 2395-2399.

ICAS Documentation, CAPEC internal report. (2001). Technical University of Denmark, Department of Chemical and Engineering. http://www.capec.kt.dtu.dk/.

Ihmels 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.

Joback K. G. and Reid R. C. (1987). Estimation of pure-component properties from group-contributions. *Chemical Engineering Communications.* 57(1-6), 233-243.

Joback K. G. and Stephanopoulos G. (1989). Designing molecules possessing desired physical property values. In *Proceedings FOCAPD* 89. 363-387.

Joback K. G. and Stephanopoulos G. (1994). Solvent substitution for pollution prevention. *AIChE Symposium Series*. 40, 98-104.

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., Chen K., Wang Y., Karelson M., Lucic B., Trinajstic N., Suzuki T. and Schüürmann G. (2000). Prediction of liquid viscosity for organic compounds by a quantitative structure-property relationship. *Journal of Physical Organic Chemistry*. 13, 80-86.

Kier L. B. and Hall H. L. (1998). *Molecular Connectivity in Structure Activity Analysis*, John Wiley & Sons, New York, NY.

Kier L. B. and Hall L. H. (2000). Intermolecular accessibility: the meaning of molecular connectivity. *Journal of Chemical Information and Computer Science.* 40(3), 792-795.

Kier L. B. and Hall L. H. (2001). Issues in Representation of Molecular Structure the Development of Molecular Connectivity. *Journal of Molecular Graphics and Modeling*. 20(1), 4-15.

Kinart C. M., Ćwiklińska A., Kinart W. J. and Bald A. (2006). Viscosities of some *n*-alkoxyethanols with ethyl tert-butyl ether at T = (293.15, 298.15 and 303.15) K. *Journal of Molecular Liquids*. 126, 135-139.

Klein J. A., Wu D. T. and Gani R. (1992). Computer aided mixture design with specified property constraints. *Computers and Chemical Engineering*. 16(5), 229-236.

Knapstad B., Skjalsvik P. A. and Øye H. A. (1989). Viscosity of pure hydrocarbons. *Journal of Chemical Engineering Data*. 34, 37-43.

Kontogeorgis G. M., Fredenslund Aa. and Tassios D. P. (1993). Simple activity coefficient model for the prediction of solvent activities in polymer solutions. *Industrial and Engineering Chemistry Research.* 32(2), 362-372.

Kuhn T. S. (1996). *The structure of scientific revolutions*. University of Chicago Press, Chicago, IL.

Lafuente C., Artigas H., Pardo J., Royo E. M. and Urieta J. S. (1996). Viscosities of 1-chlorobutane and 1,4-dichlorobutane with isomeric butanols at 25 and 40°C. *Journal of Solution Chemistry*. 25(3), 303-313.

Larsen B. L. (1986). *Prediction of phase equilibria and heat effects of mixing with a modified UNIFAC model*. Ph.D. Dissertation, Technical University of Denmark, Denmark.

Li J., Mundhwa M., Tontiwachwuthikul P. and Henni A. (2007). Volumetric properties, viscosities, and refractive indices for aqueous 2-(methylamino)ethanol

solutions from (298.15 to 343.15) K. *Journal of Chemical Engineering Data*. 52, 560-565.

Liaw H.-J., Lee Y.-H., Tang C.-L., Hsu H.-H. and Liu J.-H. (2002). A mathematical model for predicting the flash point of binary solutions. *Journal of Loss Prevention in the Process Industries*. 15, 429-438.

Macchietto S., Odele O. and Omatsome O. (1990). Design of optimal solvents for liquid-liquid extraction and gas absorption processes. *Transaction of IChemE*. 68, 429-433.

Magnussen T., Rasmussen P. and Fredenslun Aa. (1981). UNIFAC parameter table for prediction of liquid-liquid equilibria. *Industrial and Engineering Process Design and Development.* 20, 331-339.

Malagoni R. A. and Franco Jr. M. R. (2007). A simple apparatus for data solubility determination. *Fluid Phase Equilibria*. 255, 93-97.

Maranas C. D. (1996). Optimal Computer-Aided Molecular Design: A Polymer Design Case Study. *Industrial and Engineering Chemistry Research*. 35, 3403-3414.

Marcoulaki E. C. and Kokossis A. C. (1998). Molecular design synthesis using stochastic optimization as a tool for scoping and screening. *Computers and Chemical Engineering*. 22, S11-S18.

Marrero J. and Gani R. (2001). Group contribution based estimation of pure component properties. *Fluid Phase Equilibria*. 183-184, 183-208.

Martin S., Pardo J. and Urieta J. S. (2001). Viscosities for binary mixtures of 1-bromobutane and 1,4-dibromobutane with isomeric butanols at 298.15 and 313.15 K. *International Journal of Thermophysics.* 22(6), 1629-1642.

Martin T. M. and Young D. M. (2001). Prediction of the acute toxicity (96-h LC_{50}) of organic compounds to the fathead minnow (Pimephales promelas) using a group contribution method. *Chemical Research in Toxicology*. 14, 1378-1385.

Michelsen M. L. (1982). The isothermal flash problem. Part I. Stability, *Fluid Phase Equilibria.* 9, 1-19.

Michelsen M. L. and Mollerup J. M. (2007). *Thermodynamic models: fundamentals and computational aspects*. Tie-line publications, Denmark.

Miller R. W. (1996). *Flow measurement engineering handbook*. McGraw-Hill Professional, New York, NY.

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 (15), 5234-5242.

Modi A., Aumond J. P., Mavrovouniotis M. and Stephanopoulos G. (1996). Rapid plant-wide screening of solvents for batch processes. *Computers and Chemical Engineering*. 20, S375-S380.

Moggridge G. D. and Cussler E. L. (2000). An introduction to chemical product design. *Transaction IChemE-Part A*. 78, 5-11.

Morales-Rodriguez R. (2009). *Computer-aided multiscale modeling for chemical product-process design*. Ph.D. Dissertation, Technical University of Denmark, Denmark.

Morawsky N. and Martino G. T. (1997). Low VOC hair sprays with improved spray characteristics. *US Patent* 5599524.

Mundhwa M., Alam R. and Henni A. (2006). Volumetric properties, viscosities, and refractive indices for aqueous 2-((2-aminoethyl)amino)ethanol solutions from (298.15 to 343.15) K. *Journal of Chemical Engineering Data*. 51, 1268-1273.

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.

Odele O. and Macchietto S. (1993). Computer-aided molecular design: a novel method for optimal solvent selection. *Fluid Phase Equilibria*. 82(1), 47-54.

Orrick C. and Erbar J. H. (1979). Personal communication.

Pang F.-M., Seng C.-E., Teng T.-T. and Ibrahim M. H. (2007). Densities and viscosities of aqueous solutions of 1-propanol and 2-propanol at temperatures from 293.15 K to 333.15 K. *Journal of Molecular Liquids.* 136, 71-78.

Pretel E. J., Lopez P. A., Bottini S. B. and Brianole E. A. (1994). Computer-aided design of solvents for separation processes. *AIChE Journal*. 40, 1349-1360.

Qui H., Jun H. W., Dzimianski M. and McCall J. (1997). Reduced transdermal absorption of N,N-diethyl-m-toluamide from a new topical insect repellent formulation. *Pharmaceutical Development and Technology.* 2(1), 33-42.

Ralph M., DiGiulio R. M., Lee R.-J., Schaeffer S. T., Brasher L. L. and Teja A. S. (1992). Densities and viscosities of the ethanolamines. *Journal of Chemical Engineering Data*. 37, 239-242.

Raman V. S. and Maranas C. D. (1998). Optimization in product design with properties correlated with topological indices. *Computers and Chemical Engineering*. 22(6), 747-763.

Renon H. and Prausnitz J. M. (1968). Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE Journal.* 14(1), 135-144.

Ripple D. and Defibaugh D. (1997). Viscosity of the saturated liquid phase of three fluorinated ethanes: R152a, R143a, and R125. *Journal of Chemical Engineering Data*. 42, 360-364.

Rodríguez S. (1997). Densities and viscosities of binary mixtures of some cyclic ethers+chlorocyclohexane at 298.15 and 313.15 K. *Journal of Chemical Engineering Data*. 42, 1285-1289.

Roy M. N., Sarkar B. K. and Chanda R. (2007). Viscosity, density, and speed of sound for the binary mixtures of formamide with 2-methoxyethanol, acetophenone, acetonitrile, 1,2-dimethoxyethane, and dimethylsulfoxide at different temperatures. *Journal of Chemical Engineering Data*. 52, 1630-1637.

Saraiva A. (1995). Liquid-Liquid equilibria in polymer systems: theoretical and experimental study. PhD Dissertation, Technical University of Denmark, Denmark.

Saraiva A., Pleuss M., Persson O. and Fredenslund Aa. (1995). A study of the miscibility/immiscibility phenomena in low molecular weight Poly(n-butylmethacrylate)/single solvent systems by thermo optical analysis by microscopy. *IVC-SEP Internal report.*

Sastri S. R. S. and Rao K. K. (1992). A new group contribution method for predicting 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.

Satyanarayana K. C. (2009). *Computer-aided polymer design using multiscale modeling*. Ph.D. Dissertation, Technical University of Denmark, Denmark.

Sayegh S. G. and Vera J. H. (1980). Lattice-model expressions for the combinatorial entropy of liquid mixtures: a critical discussion. *The Chemical Engineering Journal*. 19, 1-10.

Shah S. M. and Fernandez C. A. (1994). Hair spray. U.S. Patent 5294437.

Shanti S. and Clifford K. S. (1993). A survey of surfactants in coatings technology. *Progress in Organic Coatings*. 23, l-22.

Shernov S. (1991). Pressurized hair spray composition. *US Patent* 5053218.

Siddhaye S., Camarda K., Southard M. and Topp E. (2004). Pharmaceutical product design using combinatorial optimization. *Computers and Chemical Engineering*. 28, 425-434.

Sinha M., Achenie L. E. K. and Gani R. (2003). Blanket wash solvent blend design using interval analysis. *Industrial and Engineering Chemistry Research*. 42, 516-527.

Smith J. M., Van Ness H. C. and Abbott M. M. (2001). *Introduction to chemical engineering thermodynamics.* Mc Graw Hill, USA.

Solvason C. C., Chemmangattuvalappil N. G. and Eden M. R. (2009). A systematic method for integrating product attributes within molecular synthesis. *Computers and Chemical Engineering*. 33, 977-991.

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(3), 782-786.

Sørensen J. M. and Arlt W. (1979). *Liquid-liquid equilibrium data collection part. 1, 2*. DECHEMA Chemistry Data Series 5, 1-3. DECHEMA, Frankfurt, Germany.

Tarng M.-R., Minamyer M., Pham A., Brownell S., Pham A., Alexander A., Shah D., Nguyen K. L., Pham M. L. and Maxey S. (2010). Exterior paint formulation. *US Patent* 7642309.

Vaidyanathan R. and El-Halwagi M. (1994). Computer-aided design of high performance polymers. *Journal of Elastomers and Plastics*. 26, 277-293.

Van der Walle G. A. M., Buisman G. J.H., Weusthuis R. A. and Gerrit E. (1999). Development of environmentally friendly coatings and paints using medium-chain-length poly(3-hydroxyalkanoates) as the polymer binder. *International Journal of Biological Macromolecules*. 25, 123-128.

Van Velzen D., Cardozo R. L. and Langenkamp H. (1972). A liquid viscosity-temperature-chemical constitution relation for organic compounds. *Industrial and Engineering Chemistry Fundamental.* 11(1), 20-25.

Van Wesenbeeck I., Driver J. and Ross J. (2008). Relationship between the evaporation rate and vapour pressure of moderately and highly volatile chemicals. *Bullettin of Environmental Contamination and Toxicology*. 80, 315–318.

Varco J. and Williams C. E. (1986). Stabilized hair spray composition and process. *US Patent* 4567040.

Venkatasubramanian V., Chan K. and J. M. Caruthers. (1994). Computer-aided molecular design using genetic algorithms. *Computers and Chemical Engineering*. 18, 833-844.

Venkatasubramanian V., Chan K. and Caruthers J. M. (1995). Genetic algorithmic approach for computer-aided molecular design. *Computer Aided Molecular Design, ACS Symposium Series*. 589, 396-414.

Villadsen J. (1997). Putting structure into chemical engineering: proceeding of an industry/university conference. *Chemical Engineering Science*. 52(17), 2857-2864.

Viswanath D. S. and Natarajan G. (1989). *Data book on the viscosity of liquids*. Hemisphere Publishing Corporation, New York.

Wang Z.-F., Wang L.-S. and Fan T.-B. (2007). Densities and viscosities of ternary mixtures of heptane, octane, nonane, and hexyl benzene from 293.15 K to 313.15 K. *Journal of Chemical Engineering Data*. 52, 1866-1871.

Weast R. C. (1984). *Handbook of Chemistry and Physics*. CRC Press, Boca Raton, FL.

Wei I.-C. and Rowley R. L. (1984). Binary liquid mixture viscosities and densities. *Journal of Chemical Engineering Data*. 29, 332-335.

Wei J. (1996). A century of changing paradigms in chemical engineering. *Chemical Technology*. 26(5), 16-18.

Weng W.-L. and Chen J.-T. (2004). Density and viscosity measurement of *n*-butylamine with hexyl alcohol isomer binary systems. *Journal of Chemical Engineering Data*. 49, 1748-1751.

Westerberg A. W. and Subrahmanian E. (2000). Product design. *Computers and chemical engineering*. 24, 959-966.

Wibawa G., Hatano R., Sato Y., Takishima S. and Masuoka H. J. (2002). Solubilities of 11 polar organic solvents in four polymers using the piezoelectric-quartz sorption method. *Journal of Chemical Engineering Data*. 47, 1022-1029.

Wibowo C. and Ng K. M. (2001). Product-oriented process synthesis and development: cream and pastes. *AIChE Journal*. 48, 1212-1230.

Wibowo C. and Ng K. M. (2002). Product-centered processing: manufacture of chemical-based consumer products. *AIChE Journal*. 48, 1212-1230.

Williamson K. J. (1994). *Macroscale and Microscale Organic Experiments*. D. C. Heath, Lexington, MA.

Wintermantel K. (1999). Process and product engineering - Achievements, present and future challenges. *Transaction of IChemE-Part A*. 77, 175-188.

Woodrow J. E., Seiber J. N. and Dary C. (2001). Predicting pesticide emissions and downwind concentrations using correlations with estimated vapor pressures. *Journal of Agricultural and Food Chemistry.* 49, 3841-3847.

Wu D.T. (1986). Prediction of viscosity of liquid mixtures by a group contribution method. *Fluid Phase Equilibria.* 30, 149-156.

Yang T., Xia S., Song S., Fu X. and Ma P. (2007). Densities and viscosities of *n*-formylmorpholine (NFM)+*p*-xylene, +*o*-xylene, +*m*-xylene at different temperatures and atmospheric pressure. *Journal of Chemical Engineering Data*. 52, 2062-266.

Yaws C. L. (1999). *Chemical Properties Handbook*. McGraw-Hill, New York, NY.

This PhD-project was carried out at CAPEC, the Computer Aided Product-Process Engineering Center.

CAPEC is committed to research, to work in close collaboration with industry and to participate in educational activities. The research objectives of CAPEC are to develop computer-aided systems for product/process simulation, design, analysis and control/operation for chemical, petrochemical, pharmaceutical and biochemical industries. The dissemination of the research results of CAPEC is carried out in terms of computational tools, technology and application. Under computational tools, CAPEC is involved with mathematical models, numerical solvers, process/operation mathematical models, process simulators, process/product synthesis/design toolbox, control toolbox, databases and many more. Under technology, CAPEC is involved with development of methodologies for synthesis/design of processes and products, analysis, control and operation of processes, strategies for modelling and simulation, solvent and chemical selection and design, pollution prevention and many more. Under application, CAPEC is actively involved with developing industrial case studies, tutorial case studies for education and training, technology transfer studies together with industrial companies, consulting and many more. Further information about CAPEC can be found at www.capec.kt.dtu.dk.

Department of Chemical
and Biochemical Engineering

DTU Building 229
Søltofts Plads
DK-2800 Kgs. Lyngby
www.kt.dtu.dk