UNIVERSITAT ROVIRA I VIRGILI

CONTRIBUTION TO THE DEVELOPMENT OF MORE SUSTAINABLE PROCESS INDUSTRIES UNDER UNCERTAINTY

Nagore Sabio Arteaga

CONTRIBUTION TO THE DEVELOPMENT OF MORE SUSTAINABLE PROCESS INDUSTRIES UNDER UNCERTAINTY

by

Nagore Sabio Arteaga

Doctoral Thesis

UNIVERSITAT ROVIRA I VIRGILI

CONTRIBUTION TO THE DEVELOPMENT OF MORE SUSTAINABLE PROCESS INDUSTRIES UNDER UNCERTAINTY

Nagore Sabio Arteaga

Doctoral Thesis

Supervised by

Dr. Gonzalo Guillén-Gosálbez
Dr. Laureano Jiménez Esteller
Dr. John T. Farrell†

†*Corporate Strategic Research, ExxonMobil Research and Engineering Company, New Jersey, US (now Vehicle Technologies, National Renewable Energy Laboratory, Colorado, US)*

Department of Chemical Engineering
SUSCAPE Research Group

Tarragona
January 2016

Universitat Rovira i Virgili

Universitat Rovira i Virgili
Department of Chemical Engineering
Av. Països Catalans, 26
47007-Tarragona, Spain
Phone +34 977 55 86 43

Dr. Gonzalo Guillén Gosálbez, Dr. Laureano Jiménez Esteller and Dr. John T. Farrell

CERTIFY:

That the present study entitled "Contribution to the development of more sustainable process industries under uncertainty", presented by Nagore Sabio Arteaga for the award of the degree of Doctor, has been carried out under our supervision at the Chemical Engineering Department of the University Rovira i Virgili.

Tarragona, 11 January 2016

Dr. Gonzalo Guillén Gosálbez   Dr. John T. Farrel   Dr. Laureano Jiménez Esteller

To my mother

ACKNOWLEDGEMENTS

This thesis would have never been possible without the knowledge, devotion and new visions that Gonzalo Guillén-Gosálbez brought to University Rovira i Virgili (URV), luckily for me in 2008, when I was finishing my Master's Degree in Chemical and Process Engineering.

That was the time when I was doing my final project degree, working in collaboration with Clariant Ibérica S.A. on the optimization and viability study of a combined heat and power plant for the site. The moment of my project defence turned out to be a very important and also a slightly disturbing one, since it implied the end of a "studies era" and the beginning of a job search in the middle of an economic crisis in a country with rampant unemployment rates.

At that specific point, two of my most admired professors during the studies at URV, Josep Font Capafons and Dieter Boer, approached me independently and suggested that research would allow me to continue with a 10 A.M. schedule while studying extremely interesting problems. By that point I already knew that optimization had the capability to happily entertain my brain during considerably large periods of time and that waking up at 10 A.M. was much more desirable to me at that time than an early morning schedule.

So I followed Dieter to the office of Laureano Jiménez, the director of SUSCAPE research group where Gonzalo was also working, to have a chat. In few days I applied for the Doctoral Programme in Chemical, Environmental and Process Engineering. Looking back, I feel extremely proud of my years at that university. The words I could use here would never cover the amount of people, moments and thanks that I would like to mention.

For instance, being a PhD student at SUSCAPE gave me opportunity to work in collaboration with ExxonMobil Research and Engineering Company, a worldwide leading process industry that not only develops the most cutting-edge process technology in the world, but also has the most brilliant people I have ever met. Amongst them was my, now, thesis director John T. Farrell.

For these, and many other reasons, I would like to first thank for their guidance, support and steadfast patience to my thesis directors, Gonzalo Guillén-Gosálbez, Laureano Jiménez Esteller and John T. Farrel. Without all of you this thesis would have never come to light.

Special thanks also to Ramkumar Karuppiah, Venkatesh Vasudevan and Nicolas Sawaya, with whom I had the opportunity to work very closely during my two international research stays at ExxonMobil. All their advices, discussions and infinite knowledge are reflected in Chapters 4 and 5 of this thesis. Thanks also to Walter Weissman, Ian Laurenzi and John Robbins, whose fruitful conversations led to ideas that are reflected throughout this work.

I lack the words to express my gratitude to Nilay Shah, my european research stay supervisor at Imperial College London, whose energy and personality have always permeated through me and continue to be a constant source of inspiration in my research career.

I will always be indebted to Neil Strachan, my current supervisor at University College London (UCL), for all his tenacious support, kindness and advice during the countless times I attempted to finish this thesis.

I wish to thank also all my SUSCAPE group colleagues, URV university folks and professors, which were once my little family in Tarragona and whom I miss every day. Special thanks are reserved to J.M. Mateo Sanz, for his wonderful statistics classes and tips, and to Núria Juanpere for her constant dedication, true kindness and invaluable help with all the administrative processes that this work involved.

And last but not least, I wish to thank my mother, because without her true unconditional and endless love this work and even myself would have never been possible.

London, 9th January 2016

Disclaimer

The views expressed here are those of the author only and do not necessarily represent the views, positions or opinions - expressed or implied - of my employers, supervisors or anyone else.

The quotes used at the beginning of each chapter have an inspirational purpose related exclusively to the concepts expressed in them, and no further relationship or connection with their corresponding authors or historical circumstances is aimed to be drawn.

SUMMARY

Over the past decades, the challenges originated as a result of high energy prices and the growing pressure to reduce greenhouse gas emissions have fuelled a large interest in energy and process systems related research. On the one hand, process industries are faced with the need to cover the increasing demand for energy as developing nations grow and developed countries continue to progress in an increasingly uncertain marketplace, and on the other hand, the resources that have traditionally supported this continued progress begin to show environmental impacts that could threaten the sustainable development of species in the world. As a consequence, the present situation could be described as driven along three main edges: energy, sustainability and uncertainty.

Of particular relevance for these problems is research on computer-aided systems technology to develop strategies for investigating the impact of process industries on both, the system efficiency and its life cycle environmental impact. In this sense, Process Systems Engineering (PSE) offers a unique set of tools that are capable of applying traditional engineering and scientific knowledge to systemic problems, thereby enlarging the scope of traditional chemical engineering to larger system scales while allowing the application of robust and systematic tools to more complex systems problems. Hence, the new emphasis on energy and sustainability experienced in the area has been appended to its other more traditional computational and uncertainty related problems.

In this sense, the general goal of this thesis is to explicitly address these challenges by first making a step towards closing the gap between science-based and systems-based research in PSE. This problem is addressed through the integration of techniques and theories from different disciplines into advanced mathematical programming programming frameworks able to deal with both, the environmental and the uncertainty challenges in the design and planning of more sustainable process industries. For this purpose, multi-objective optimization is proposed as the core mathematical programming framework able to represent the effects of these, sometimes, conflicting criteria in the design of process systems.

In particular, a set of multi-objective optimization tools able to deal with both, uncertainty sources and the life cycle environmental impact, are proposed for two problems of process design. Thus, the first half of the thesis is devoted to the design and planning of hydrogen supply chains and the second half to the design of a large-scale complex industrial process plant. The problem of hydrogen infrastructure design has been mentioned in the scientific literature to be of paramount relevance for enabling the development of hydrogen as an energy vector with the potential to drive the transition towards a more sustainable energy system, whereas the problem of whole industrial process optimization has been a traditionally challenging one in PSE particularly due to the computational complexity involved in accurately representing process unit operations. In addition, a strategy for reducing the number of redundant objectives in cases where more than one environmental life cycle assessment (LCA) metrics need to be explored is also presented.

The energy challenge is addressed, first by providing two frameworks for designing hydrogen infrastructures, one capable of mitigating the effects of uncertainty in energy prices and another one able to optimize the economic performance and any life cycle environmental metric defined by standard LCA methodologies, identifying robust and non-redundant more sustainable hydrogen supply chain designs. Hydrogen presents several advantages as an energy vector, that are mainly given by its potential to become environmentally friendy and by its adaptability to the current energy system conditions, where it can play roles ranging from energy storage vector for the electricity system, to being used as a fuel or chemical agent in industrial operations.

With regards to sustainability, LCA has recently emerged as a key element for environmental impact assessment that allows to trace emissions and waste generated in industrial processing activities from “cradle-to-grave” in a holistic manner. Therefore, our approach is to append the LCA metrics as additional criteria to be optimized. Combining multi-objective optimization with LCA allows for the automation of the search for process design solutions that can be more environmentally friendy. By proposing a solution based on principal component analysis that can identify redundant environmental metrics, we ensure that all the metrics of interest can be explored, while decision-makers can be aware of particular interactions beteween them, therefore making the problem analysis and decision steps more tractable.

Uncertainty is addressed by formulating stochastic mathematical programming frameworks, that allow to quantify and evaluate the effects on parameter uncertainties at the design step. By appending risk metrics as additional criteria to be optimized, these models are able to represent different attitudes that decision makers can exhibit towards the risk associated to changing conditions. Thus, the mathematical models provided are able to account for the methodological uncertainty associated to classical deterministic frameworks. In addition, appending LCA metrics as additional criteria to be optimized allows to account for the methodological uncertainty associated to traditional frameworks that only dealt with economic performance in a similar way. Furthermore, by focusing on the uncertainty and evaluating the capabilities of the modelling frameworks in different conditions and with different structural characteristics, more robust models can be provided that are able to account for the many sources of uncertainty that usually affect modelling tasks and in turn the associated decision making.

Thus, the second chapter of this thesis is opened with a problem for hydrogen supply chains design that allows to control for the variability associated to uncertain in energy prices (i.e., operating costs). The mathematical model is formulated as a multi-objective multi-period stochastic programming MILP problem that is capable of simultaneously optimizing the expected economic performance and of controlling for undesirable outcomes related to volatile energy prices. This latter task is accomplished by appending a the worst-case value, a financial risk metric, to the economic performance as an additional criterion to be optimized. In addition a two-step sequential algorithm is presented, which is able to expedite the search of the corresponding Pareto set by one order of magnitude. The results in this case showed that, in the actual conditions, a hydrogen network constituted by centralized steam methane reforming plants would be the most economic solution, whereas coal gasification would be a more robust design more able to cope with the uncertainty energy prices.

Next, the hydrogen design problem is reformulated to its deterministic form in order to solve high dimensional multi-objective model, where the economic performance of the network design is simultaneously optimized with eight different life cycle environmental metrics. Note that these metrics are used as part of the well-known Eco-Indicator 99 life cycle assessment (LCA) methodology for calculating 4 damage categories. The corresponding pairwise Pareto sets are represented, and three main hydrogen supply chain design trends are identified for them. The problem scope is enlarged by applying a principal components analysis (PCA) to the post-optimal analysis of the environmental metrics in order to determine the conflicting and redundant objectives. The results show that only four of the eight initial metrics would suffice for representing the problem with a very large extent (measured by the variance resulting from the PCA), and the representative network designs that minimize the conflicting environmental impact metrics showed decentralised hydrogen supply chains of compressed hydrogen from steam methane reforming and wind electrolysis.

Then we proceed to present a novel formulation for the optimal design and operation of single-site large-scale industrial process plant. The problem is posed as a deterministic multi-objective MINLP problem that is capable of simultaneously optimizing the economic performance and three different environmental impact metrics. Several model tests where performed in order to understand the impacts of different uncertainty sources on it. As a result, the model was tested with two different energy price datasets, for the bi-criterion case including only one environmental impact metric and for the multiple criteria case of the three environmental metrics and the economic objective, the demand constraint was tested for a fixed and a flexible case and the product quality constraints that limited the process operation were also tested by means of a sensitivity analysis. The results showed that the economic performance of the model exhibited more flexibility than its environmental performance.

Finally, with the aim of considering the uncertainties that tend to affect life cycle inventory of emissions associated to LCA, the deterministic problem presented in the previous chapter was reformulated to its environmental stochastic counterpart. In order to allow for the representation of different attitudes that decision-makers may exhibit towards environmental risk, we appended to the economic objective function three different stochastic and risk metrics represented by the expected environmental performance, the worst-case value and the downside risk. The underlying multi-objective formulation capable of simultaneously optimizing for all the above metrics is able to represent any type of probability distributions of the uncertain parameters with or without correlation. Without loss of generality, the problem was solve for three different correlation cases and lognormal probability distributions, which tend to be a more common case in LCA.

The presented set of multi-objective tools are devised as particularly suitable to deal with the challenges that motivated the work, as first, they are able to design process systems, that by the consideration of life cycle assessment metrics have the ability to become more sustainable, second, they can be applied to single-site and multi-site industrial process and energy systems, and third, they are specially suited to produce robust and systematic solutions capable of addressing the three major uncertainty sources affecting systems problems: parameter, model and methodological uncertainty.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Generic abbreviations

CAPE	Computer aided process engineering
CDF	Cumulative distribution function
CPU	Central processing unit
GAMS	General algebraic modeling system
GHG	Greenhouse gas
IDE	Integrated development environment
LCA	Life cycle assessment
LCI	Life cycle inventory
LCIA	Life cycle impact assessment
lnN	Lognormal distribution
MILP	Mixed integer linear programming
MINLP	Mixed integer nonlinear programming
moMINLP	Multi-objective mixed integer nonlinear programming
MOO	Multi-objective optimization
N	Normal (Gaussian) distribution
NL	Variable that is not at its limiting value in an optimal solution (i.e., the constraint enforcing it is active and its Lagrangean multiplier is zero)

NLP	Nonlinear programming
Limits	Variable that is at its limiting value in an optimal solution (i.e., the constraint enforcing it is active and its Lagrangean multiplier is not zero)
PC	Principal component
PCA	Principal component analysis
PSE	Process systems engineering
SC	Supply chain
SCM	Supply chain management

Process abbreviations

BG	Biomass gasification
CCS	Carbon capture and storage
CG	Coal gasification
C1-C7	Emission compounds
C1-C7	Emission compounds
E	Electrolisis
Metric 1-3	Environmental impact performance metrics used
M	Hydrogen supply chain multi-objective MILP model for total discounted cost and 8 different environmentaal metrics optimization
MOP	Hydrogen supply chain multi-objective stochastic MILP model for expected total discounted cost and worst case optimization
$MOPN_{DR}$	Industrial system multi-objective stochastic MINLP model for downside risk and economic performance optimization
$MOPN_{WC}$	Industrial system multi-objective stochastic MINLP model for worst case and economic performance optimization
MP	Master LP model for P solution strategy
P	Hydrogen supply chain single-objective stochastic reformulated MILP model for (MOP) solution strategy
PN	Industrial system multi-objective MINLP model for environmental and economic performance optimization
PNA	Industrial system single-objective reformulated MINLP model for (PN) solution strategy
Prop A-F	Product quality specifications
P1-P4	Final products
P5-P7	Byproducts
R1-R4	Raw materials
SMR	Steam methane reforming
SP	Slave MILP model for (P) solution strategy
T1-T4	Final products storage tanks
T5-T7	Byproducts storage tanks
U1-U4	Process utilities - fuel
U5-U7	Process utilities - steam, electricity and cooling water

Unit 1-9	Process units
WE	Wind electrolisis

Commercial abbreviations

DICOPT	Discrete and continuous optimizer (MINLP solver)
AlphaECP	Extended cutting plane MINLP solver
SBB	Standard branch & bound MINLP solver
NPV	Net present value
SAA	Sample average approximation

CHAPTER 1

INTRODUCTION

If the Lottery is an intensification of chance, a periodic infusion of chaos into the cosmos, then is it not appropriate that chance intervene in ***every aspect*** *of the drawing, not just one?*

Jorge Luis Borges, The Lottery in Babylon, 1941

A new emphasis on energy and sustainability has emerged recently in the area of Process Systems Engineering (PSE). It arises as a response to the observed and projected continued appetite for energy in the world and the consequently increasing carbon emissions related to the combustion and use of fossil resources (IEA, 2015; Grossmann and Guillén-Gosálbez, 2010). The underlying challenges of how to supply that energy and how to reduce the associated emissions, lie within the basis of chemical engineering core expertise, and given the enlarged scope of PSE to cover larger spatial and temporal scales of systems (see Figure 1.1), these can be posed as complex whole systems problems where the largest unit in scale to be studied can be seen as the earth and its surrounding environment. Needless to say is the fact that the nature of these problems calls for an interdisciplinary perspective, where the cross-fertilization of different areas of knowledge appears as an invaluable tool for fostering societal advances.

This thesis explicitly attends the demand of these challenges by making a step towards closing the gap between science-based and systems-based research in the area of PSE, stated first by Grossmann and Westerberg (2000) at the beginning of this century. Through the integration of techniques and theory from different areas of knowledge, the main goal of this work is to develop advanced mathematical programming tools for the synthesis and planning of more sustainable process systems[1].Ultimately, the goal of the products of this thesis is to serve as decision-support tools to inform at the different steps of the decision-making processes involved on energy and sustainability problems, specifically when these take place in the presence of uncertainty.

More specifically, knowledge from different fields is integrated within a unified framework in this thesis: from management science through the use of the supply chain concept to solve large scale energy systems problems and risk management to control for uncertain undesirable outcomes, from sustainability with the implementation of life cycle assessment (LCA) as a central piece within the frameworks and from statistics by applying principal component analysis (PCA) for addressing the problem of objective visualization and decision-making in the underlying multi-objective optimization (MOO) problems. As a core element, computer-aided process engineering (CAPE) multi-objective mathematical programming tools are used for formulating general modelling frameworks able to deal both, with uncertainty and sustainability goals. These are tested through their application to two specific synthesis problems: a multi-site problem on the design and planning of hydrogen supply chains, and a single-site problem on the synthesis and operation of complex industrial chemical plants.

The importance of uncertainty on optimization has been acknowledged since the beginnings of mathematical programming history (Dantzig, 1955), with the problem of planning under uncertainty still being considered as one of the most important open problems in optimization (Sahinidis, 2004). Different classifications of uncertainty have also been proposed to best tackle the problems arising from the need to assess and control the uncertainty governing some of the model parameter spaces. Nevertheless, to the extent of our knowledge, in the area of process systems engineering these classifications have traditionally been focused on the uncertainty that affects the parameters, somehow leaving aside the modelling activity the structural and methodological uncertainties associated to the framework itself and the choices, perspectives or objectives adopted in it for decision-making.

Although all the aforementioned types of uncertainty have a key role and are addressed in different parts of the thesis, one its contributions lies on having widened the scope of the uncertainty considered in PSE problems, by recognizing and embedding within the modelling framework the uncertainty due to choices, also known as methodological uncertainty. More specifically, the methodological uncertainty associated with classical single-objective optimization frameworks is implicitly addressed by proposing a multi-objective optimization approach as the default basis for dealing with the problems of

[1]A process system here refers to the set of activities carried out in the processing industry, where raw materials are converted into final products generally by using a set of energy vectors and/or processing equipment. By this definition, a process system includes the energy systems involved in its processing activities and at the same time it is also enclosed as part of a wider energy system

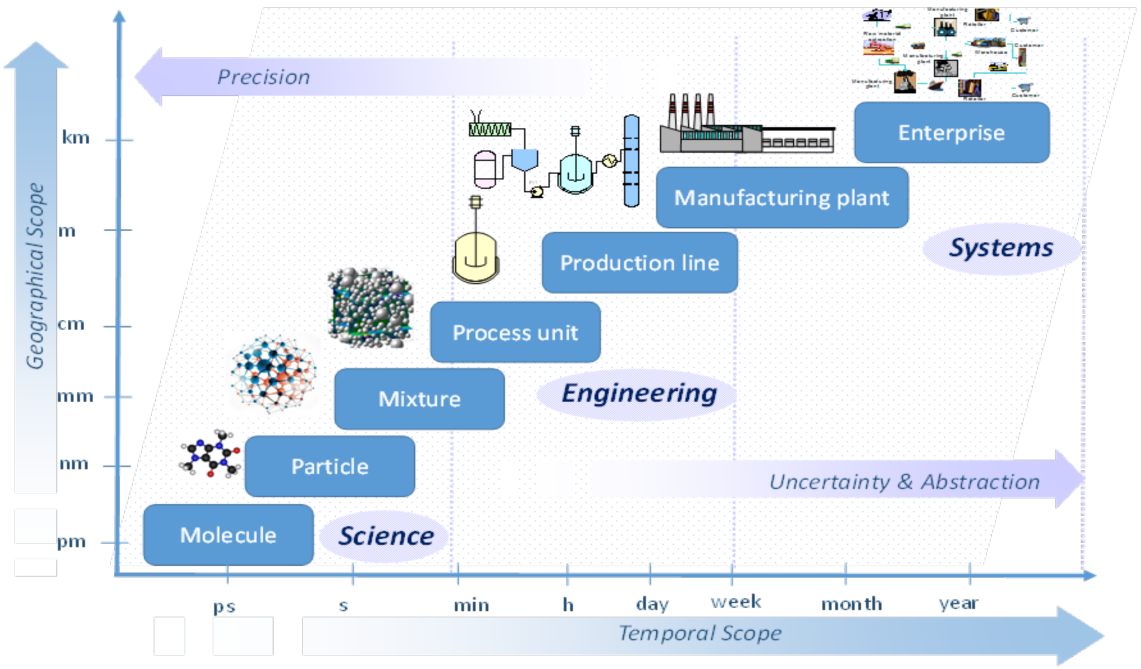


Figure 1.1: Chemical supply chain in a multidisciplinary scale perspective

design and planning of more sustainable process systems. This perspective is introduced in more detail in Subsection 1.5.3. The background, motivation and major challenges that this work aims to address are presented in Section 1.1. We next give an overview of the PSE field in Section 1.2, where the applied nature of the area is depicted and its activities are mapped according to its now familiar temporal scopes and decision levels. Section 1.3 delivers a brief introductory note to the science basis of the mathematical programming techniques used for process design and optimization. Then our approach for dealing with sustainability in process systems design is explained in Section 1.4, and as previously mentioned, Section 1.5 introduces our perspective for dealing with uncertainty in process systems optimization. Finally, in Section 1.6 a schematic outline of the present work is given, both on a conceptual and on a content basis. Note that the aim of the present chapter is mainly to serve as a roadmap for this thesis, by providing a perspective and an overview of the background, challenges and methodology used. The specific literature reviews, methodological details and model formulations are presented within each of the chapters.

1.1 Background and motivation

Over the past decades, challenges originated as a result of high energy prices and the growing pressure to reduce greenhouse gas emissions have fueled a large interest on energy and process systems related research. Of particular relevance for these problems is research on computer-aided systems technology to develop strategies for investigating the impact of the design of process industries on both, the system efficiency and its life cycle environmental impact. With regard to the environmental performance of a system, life cycle assessment has emerged as a key element that allows to trace and represent the emissions and waste related to the relevant activities associated to a product or process from “cradle-to-grave” in a holistic manner (Guinée et al., 2002).

Life cycle assessment (LCA) was originally envisaged as a descriptive tool for identifying the main sources of environmental impact of a product or process over its entire life cycle (see Subsection 1.4.1). The lack of a systematic approach in LCA to improve the environmental performance of a system has since very beginning presented a major drawback of the framework. In this sense Azapagic and Clift (1999) where the first to propose the integration of LCA and multi-objective optimization (MOO) as an effective method to overcome this limitation. As a result of these trend and its direct application to the current energy challenges, multi-objective optimization tools have started to play an increasingly important role in the design and planning of more sustainable process systems during the last decade (Grossmann and Guillén-Gosálbez, 2010; Pieragostini et al., 2012). Subsection 1.4.2 introduces the general methodology used for coupling MOO and LCA and provides the references to the corresponding chapters and sections where the specific problem formulations and solution strategies can be found.

To our knowledge, the work by Grossmann et al. (1982) constitutes the first attempt to implement environmental considerations in an MOO framework in PSE, where the local toxicological impacts of the plant and the economic performance were simultaneously optimized in a synthesis problem for industrial chemical complexes. In general nowadays, when the problem addressed involves the representation of multiple sites along large spatial and temporal scopes, the formulation takes the form of a mixed-integer linear programming (MILP) problem (Bojarski et al., 2009; Hugo and Pistikopoulos, 2005; Guillén-Gosálbez et al., 2010; Mele et al., 2011). This is because in these problems it is possible to represent the capacity limitations *via* simplified linear constraints. However, design problems at the plant level require smaller temporal and spatial resolutions within the chemical supply chain (see Figure 1.1), which lead to non-linear equations to describe the process system operations in more detail. This gives rise to more complex mixed-integer non-linear programming (MINLP) problem representations (Guillén-Gosálbez et al., 2008). While considerable advances have been accomplished in the solution and application of MILP and MINLP models over the past decades (Grossmann, 2002; Sawaya, 2006; Karuppiah and Grossmann, 2012), challenges still remain those of modelling and solving large scale synthesis and planning problems for sustainable process systems (Grossmann and Guillén-Gosálbez, 2010). This constitutes a major barrier that has so far constrained the application of these studies to the academic environment, despite their potential for improving industrial decision-making practices. Details on the general formulation of these problems is introduced in Section 1.3 and Figure 1.3.

A major limitation present in most recent optimization-based frameworks that incorporate life cycle analysis like the ones presented by Guillén-Gosálbez et al. (2010); Zamboni et al. (2009); Mele et al. (2011); Salcedo et al. (2012) is their reliance on deterministic formulations, which assume that a "base-case" scenario will be realised with perfect accuracy. As mentioned above, the already complex nature of the underlying large-scale MILP and MINLP multi-objective optimization formulations, which result in high computational burdens, together with the applied nature of the field constitute significant barriers for this type of research. However, in practice the various forms of uncertainty inherent in pricing, supply, demand, unit operation and life cycle inventories require more robust decisions that can potentially optimize the economic and environmental performance in a wide variety of scenarios.

Although there has been already considerable progress in the area of optimization under uncertainty, mainly for solving large scale complex stochastic formulations (Sahinidis, 2004; Geletu and Li, 2002), little work has been done to incorporate life cycle analysis for the purpose of driving robust and effective decisions able to mitigate uncertainty. Although uncertainty analysis is common in environmental and LCA studies is a widely studied subject (Huijbregts, 1998, 2001; Zelm et al., 2009; Boithias et al., 2016), to the best of our knowledge, Guillén-Gosálbez and Grossmann (2009, 2010) were the first to introduce a multi-objective MILP formulation that specifically addressed the uncertainty present in the life cycle inventory associated with the operation of a supply chain problem. Then, Sabio et al. (2014) formulated a large scale process MINLP problem in which the uncertainty associated to the life cycle inventory of emissions was represented using correlated distributions of different uncertainty scales.

The approach introduced in Sabio et al. (2014) was next applied to a MILP problem of petrochemical supply chains design by Reyes-Labarta et al. (2014), showing the importance of the flexibility that the process model exhibits for reaching more robust outcomes. More recently, stochastic formulations dealing mostly with uncertainty in demands and supply of MOO approaches coupled with LCA for MILP supply chain formulations have been presented by Gebreslassie et al. (2012); Balaman and Selim (2014); Osmani and Zhang (2014). In spite of all these advances, the area of optimization under uncertainty to date has only focused on investigating parametric uncertainty, somehow relying on activities or operations outside the modelling scope for dealing with uncertainties of structural or methodological nature, despite being the latter the ones having the largest impact in the solutions is exposed by Boithias et al. (2016). Section 1.5 portrays in more detail our novel perspective for dealing with uncertainty in PSE.

The problem of uncertainty characterization deserves a field of study on its own. The classification most commonly used in the areas of life cycle assessment and decision making is followed in this work. This classification distinguishes between parameter, model or structural uncertainty and uncertainty due to methodological choices. Structural uncertainty refers to the extent to which structural features of the model (*i.e.*, constraints, parameters and variables included) adequately capture the nature of the process described. And methodological uncertainty, on the other hand, arises when there are different normative views about what constitutes the "correct" approach for optimal decision-making (*i.e.*, objective functions to be included in the formulation, allocation methods) (Huijbregts, 2001; Bilcke et al., 2011). Section 1.5 extends the interpretation of this concepts and their application in this work.

Optimization under uncertainty has classically dealt with problems more closely related to the science-based research, where parameter uncertainty tends to play a central role. Parameter uncertainty arises owing to uncertainty in the choice of parameter values (Bilcke et al., 2011). Parameters used in the models are usually economic, demographical, epidemiological and experimental in nature. It is well known that some aspects related to structural uncertainty are well known to be parametrizable (Bilcke et al., 2011), Nevertheless, without undermining the importance that all types of uncertainty have, methodological uncertainty related to the perspective taken or function to be optimised represents a major challenge in optimisation based frameworks that is rarely discussed. As a consequence, challenges still remain both, those of addressing other important sources of uncertainty and those of designing whole and robust energy and process systems more sustainable.

Despite the general nature of the modelling frameworks proposed, the specific problem addressed and its scope play both significant role in PSE. The problem-specific constraints and nuances required to accurately model different process systems determine the type of formulation finally obtained, and in turn also its adequate solution strategy (see Section 1.3 and Section 1.4). In this respect, single-site problems have historically received more attention in the literature than multi-site problems (Grossmann, 2004, 2005). The case of multiple plants (chemical supply chains) has only recently emerged as an important area of research in PSE during the present decade (Grossmann and Westerberg, 2000), and as a consequence, the integration of different decision-making levels (*i.e.*, strategic, tactical, operational) still remains an important challenge in PSE (Grossmann, 2012).

The overall goal of this thesis is, then, to contribute to the development of CAPE tools for the design and planning of more sustainable processes under uncertainty. The problem is tackled by establishing new connections between the science-based fields of chemical engineering and statistics, and the systems-based fields of management science and life cycle assessment. Using basic tools from the PSE field, such as multi-objective optimization and optimization under uncertainty as a bridging connection, new theoretical and computational insights capable of exploiting the rich knowledge developed in all these areas are obtained.

In particular, the major theoretical challenges addressed in this thesis are the following.

- First, by implementing a multi-objective formulation as a basis for the synthesis and planning of more sustainable process systems, we provide a tool for analyzing the methodological uncertainty associated with classical single-objective optimization.

- Second, the combination of multi-objective optimization (MOO) and life cycle analysis (LCA) provides a systematic methodology capable of analyzing solutions that reduce the environmental impact to the design of process systems.
- Third, life cycle analysis is incorporated to optimization under uncertainty with the aim of driving more robust and effective sustainable process design options able to mitigate uncertainty.
- Fourth, risk management metrics are proposed as effective modelling tools capable of controlling undesirable outcomes, while representing different decision-maker attitudes towards risk.
- Fifth, the problems addressed in this work take into account strategic and tactical aspects, as well as single-site and multi-site process systems, therefore contributing to fill the gap in the PSE literature for frameworks integrating different spatial and temporal scales with wider systemic perspectives.

To the extent of our knowledge, this particular set of MOO tools coupled with LCA has first been brought in this work to the real corporate environment by formulating the complex industrial scale problem right from the start (see Chapter 4, Chapter 5 and Sabio et al. (2014)). Note that we present contributions on both, the modelling and algorithmic dimensions of the problem, with the main focus placed on the former aspect. Hence, in addition to developing more holistic mathematical models, we provide effective tailor-made algorithms that expedite the solution procedure.

1.1.1 Objectives

The objectives of this thesis are:

- Devise a systematic and holistic framework able to mitigate and assess the financial risk associated to volatile energy markets for the design and planning of hydrogen supply chains under uncertainty.
- Extend the problem of environmental hydrogen supply chain design to account for all damage categories and environmental impacts included in LCA methodologies, while reducing it to the smallest possible representative set.
- Formulate a novel mathematical programming framework for large-scale complex industrial process plants able to simultaneously optimize the economic performance and the life cycle environmental impacts.
- Extend this framework to systematically produce robust solutions in the presence of uncertainties in LCA.

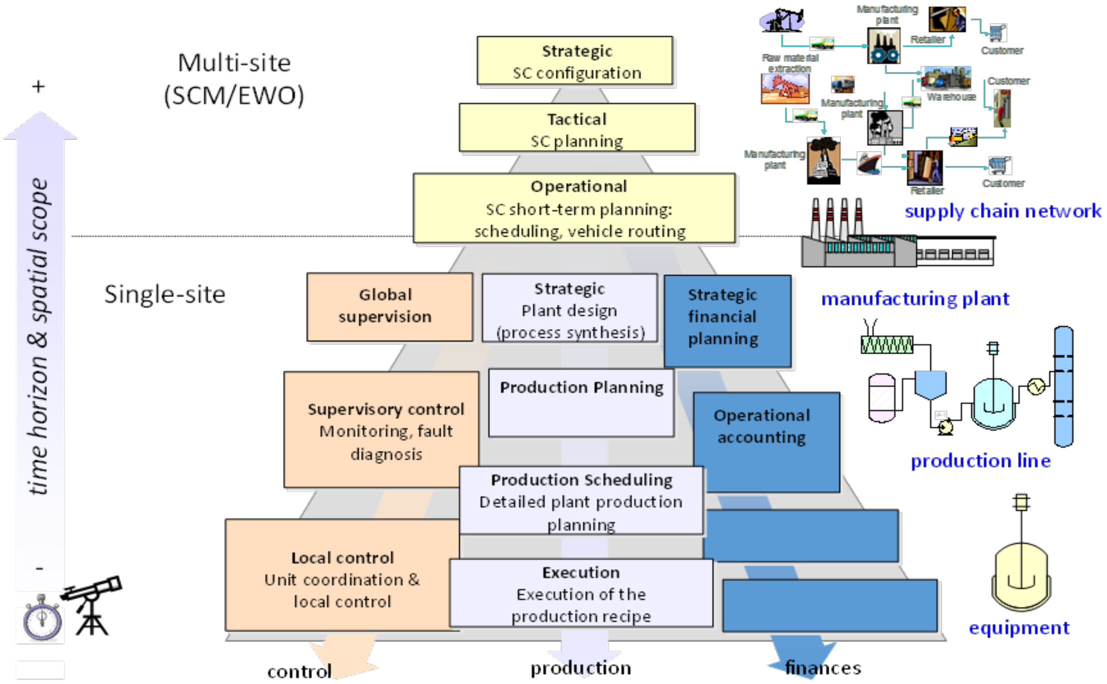


Figure 1.2: Hierarchical decision levels in PSE (Grossmann and Guillén-Gosálbez, 2010)

1.2 Process systems engineering

Process systems engineering (PSE) is the research branch of chemical engineering that aims to solve industrial problems through the development of methods and tools that effectively combine knowledge from science, engineering and systems. PSE interests range from the molecular level to the management of the product supply chain (see Figure 1.1), being the synthesis and planning of processes a central aspect of its associated research(Grossmann and Westerberg, 2000).

A common and practical approach for classifying activities within PSE according to their temporal scale distinguishes between for tactical, operational and strategic planning decisions. Strategic planning involves long-term decisions (usually 1–2 years), like the design and redesign of chemical plants. Tactical planning involves decisions on the medium term (3–6 months). These include, among others, purchases of raw materials and production levels to be set at the chemical plants. Finally, operational planning deals with activities in the short term (generally, the time horizon extends from one or more days to a week), such as the scheduling of manufacturing tasks and strategies for control and supervision at local levels. All these temporal scale levels are found at the two major geographical scale units, the single-site (*i.e.*, chemical or process plant) and the multi-site systems (*i.e.*, supply chains), as depicted in Figure 1.2.

At the top of the pyramid in Figure 1.2 we can find the entire supply chain (SC), in which the plant is immersed. This is the field of study of the new emerging area of PSE known as “Chemical Supply Chain” or “Supply Chain Management” (SCM) (Grossmann and Westerberg, 2000; Simchi-Levi et al., 2000), whose goal is to provide support tools to take decisions associated with the optimal design and management of chemical production and distribution networks. The largest scale decision units in PSE at the moment are therefore the supply chain and the manufacturing plant. These two levels have in common systems-type traits, such as the fact of being influenced by all decision levels (*i.e.*, strategic, tactical and operational) and the exposure to uncertainties whose effects are correlated to their scale (*i.e.*, large), as indicated in Figure 1.1. This constitutes one of the major reasons supporting the explicit consideration of methodological uncertainty at these decision levels, since its effects can not only be pervasive throughout the scales, but also irreversible due to their magnitudes.

1.3 Mathematical Programming

Amongst the set of tools conforming the science base of PSE, mathematical programming, also referred to as optimization, has always played a central role and has experienced notable advances in the field due to its widespread applications (Grossmann, 1989; Grossmann and Westerberg, 2000; Grossmann, 2002, 2004, 2005; Grossmann and Guillén-Gosálbez, 2010; Grossmann, 2012, 2014). In mathematical programming, optimization problems take the form of a mathematical framework MOO, as expressed in Equation 1, where an objective function set F is minimized or maximized without violating any problem-specific resource or design constraints (g, h).

$$
\begin{aligned}
(MOO) \qquad & \min_{x,y,z} \left(F_1(x,y,z), F_2(x,y,z), ..., F_n(x,y,z)\right) \\
s.t. \qquad & h(x,y,z) = 0 \\
& g(x,y,z) \leq 0 \\
& x \in \Re,\ y \in \{0,1\},\ z \in N
\end{aligned}
\tag{1}
$$

The objective function set and the mathematical constraints $(F_n], g, h)$ are formulated as algebraic equations, which can be linear or non-linear. The variables on whose domain the equations are defined (x, y, z), can take the form of continuous or discrete variables. In the discrete space, binary variables are used to represent the logic decisions and integer values to represent the number of units.

When all the algebraic equations are linear and all the variables are continuous, the model takes the form of a linear programming problem (LP). In a simple case composed by two variables, the problem could could be geometrically represented as shown in Figure 1.3 a, where the lines represent the linear equations (*i.e.*, $g(x, y, z), h(x, y, z)$) and the grey area would be the corresponding feasible region enclosed within the equality and inequalities. On the other hand, if non-linear equations are present in any of the problem equations (*i.e.*, objective function or constraints), but all the variables are still continuous (*i.e.*, type x), a non-linear programming problem is defined (NLP). As can be observed in Figure 1.3 b, the corresponding feasible region becomes geometrically more sophisticated. Finally, when discrete (*i.e.*, integer (type z) or binary (type y)) variables are present in the problem equations, the formulation takes the form of a mixed-integer linear programming problem (MILP) when all the equations are linear, and a mixed-integer non-linear programming problem (MINLP) when non-linear equations are present (see Figure 1.3 c and Figure 1.3 d respectively).

Figure 1.3 graphically shows how the presence of integer variables reduces the feasible region of the originally continuous formulation and makes it non-convex, thus increasing the complexity of the problem resolution. This classification effectively depicts the problems in increasing order of resolution complexity and helps explaining the structure of this work. The first half of the thesis, (*i.e.*, Chapter 2 and Chapter 3), concentrates on two different MILP problem formulations of a hydrogen supply chain design. Section 2.2 introduces a stochastic formulation of the problem that accounts for uncertain energy prices and Section 3.2 depicts its deterministic counterpart, where methodological uncertainty is more explicitly addressed by optimizing eight different environmental metrics agains the economic performance. Their complete formulations are presented in Section 2.3 and Section 3.3 respectively and the corresponding case studies can be found in Section 2.5 and Section 3.5.

Chapter 4 and Chapter 5 present a new large-scale MINLP framework representing a synthesis problem for a complex industrial process plant introduced in more detail in Section 4.2. The deterministic formulation of the problem is detailed in Section 4.3 and its stochastic counterpart in Section 5.3. The solution strategies are followed for these large-scale multi-objective deterministic and stochastic MINLP problems are presented in Section 4.4 and Section 5.4. Note that to this point we have so far followed the approach of solving the deterministic and stochastic programming formulations of the same problems as a systematic methodology to account for uncertainty at the design stage. The complete details of our perspective for optimization under uncertainty and the reasons for choosing stochastic programming as our basic framework to deal with parameter uncertainty are set in Section 1.5.

1.3.1 Solution approaches for multi-objective problems

Multi-objective optimization problems are formulated when a conflict between the criteria of interest is envisaged. The interest in resolving the problem lies in unveiling the trade-off relationship between both objectives, which can range from non-existent or insignificant, to not only relevant but also exploitable according to the circumstances. For this reason, when the objectives can be represented by pairs, the Pareto set or Pareto frontier (see Figure 1.8 in Subsection 1.5.3) is an efficient method for describing the

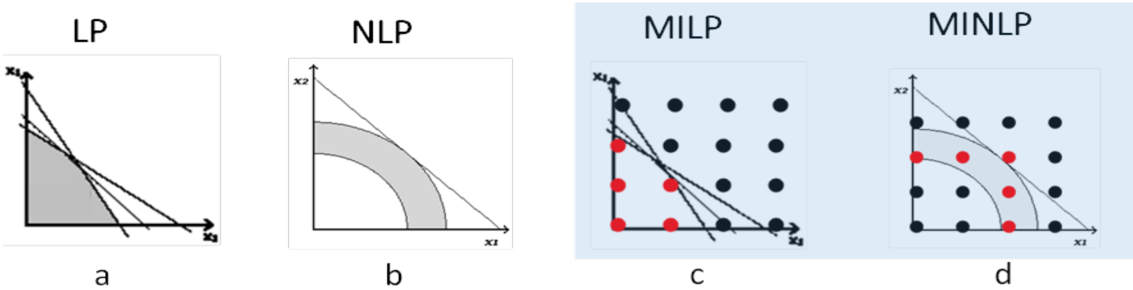


Figure 1.3: Graphic representation of feasible regions for major classes of optimization problems

relationship between the objectives of interest. Despite the advantages provided by having this additional information, multi-objective optimization problems also suffer from associated additional complexities. For instance, if the problems of interest are more ill-defined and the optimization task is approached through a more exploratory perspective, the visualization of all the conflicting relationship combinations and its analysis can turn into an impracticable problem whose resolution tends to resort to human-model iterative interaction methods to be solved (Branke et al., 2012). This subject is briefly introduced in Subsection 1.4.3, where references to the chapters and sections of this work where theory and formulations for objective reduction techniques are presented in more detail can be found.

The resolution of a multi-objective optimization problem involves finding the surface that defines the trade-off between the objectives and thus specific strategies are required for this task. The three main types of MOO solution approaches, as presented by Grossmann (2014) are:

1. Those based in the transformation of the problem into a single-objective one
2. The non-Pareto approaches which use search operators based on the objectives to be optimized
3. The Pareto approaches which directly apply the concept of dominance

Here the first class of approaches is chosen due to their suitability to be applied in conjunction with standard exact algorithms (*i.e.*, LP, NLP, MILP), as opposed to the second and third classes (Grossmann, 2014). Within the first class or single-objective multi-objective optimization methods we can find:

- Weighted-sum method or aggregation methods, based on aggregating the objectives into a single vector using pre-defined weights for each objective.
- Goal programming approach, where the objectives are assigned a pre-defined target value that need to reach and the objective function is formulated as an aggregation of the previously defined satisfaction values.

- ϵ-constraint method (Haimes et al., 1971), where one of the objectives is considered as primary and the others are posed as secondary objectives formulated as constraints *via* auxiliary ϵ parameters.

As can be observed, the first two methods imply the use of weights or "a-priori" judgements before the problem is resolved. The objective then turns into an algebraic combination of the objectives. Goal programming can be considered a particular case of the weighted-sum method, where the objectives are substituted by their corresponding satisfaction values. The resolution of an MOO problem following either of these approaches gives rise to either a single solution if single weights are used for each objective, or to a surface predefined by the weights used when several weights are used. In addition to the necessity to include "a-priori" information, as explained in Subsubsection 2.3.6.6, the weighted-sum method is only rigorous for the case of convex Pareto sets, whereas the ϵ-constraint method is also rigorous for the non-convex case, which turns to be our case in all the problems solved in this work. We therefore select the ϵ-constraint method as our default approach to solve the multi-objective MILP and MINLP formulations of this work for two major reasons, first its adequacy for solving non-convex Pareto sets, and then its capability of reflecting the underlying existing relationship between the objectives without the need to resort to "a-priori" user information.

The solution approach and its algorithmic implementation is introduced in Section 4.4 and represented in Figure 4.7 for the case of simultaneously optimizing the economic performance and three different environmental metrics. It is also applied in Subsubsection 2.3.6.6 for optimizing the expected value of the economic performance and the worst case metric, in Section 3.4 for optimizing one economic performance metric against eight different environmental criteria and in Section 5.4 as an extension of the stochastic counterpart of the problem solved in Section 4.4 where the expected economic performance is optimized against the worst case value and the downside risk metrics. The corresponding Pareto sets are illustrated in Figure 2.3, Figure 2.4, Figure 3.2 to Figure 3.9, Figure 4.8, Figure 4.12,Figure 4.13,Figure 4.16, Figure 5.5 and Figure 5.6 in their corresponding chapters.

Note that here we present the general case of the problem, where one or more objective functions can be formulated. Subsection 1.4.2 introduces the specific deterministic problem formulation $MOOe$, where one economic performance criteria is optimized against several environmental impact metrics, which is the case presented in Chapter 3. Thereafter, Subsection 1.5.1 presents the specific stochastic problem formulation $MOOr$, where the expected value of an economic performance metric is optimized against a risk metric (*i.e.*, worst case or downside risk).

1.4 Environmental impact assessment

The sustainability challenge, defined more generally as the endurance of systems and processes, is addressed here by appending the environmental impact as a necessary criterion to be optimized along with the economic performance in the strategic design and planning of process systems. Environmental impact assessment methods are evidence based procedures used to assess the environmental effects of a system. Since their formal implementation in the United States in the early 70s, they have been increasingly used around the world. Nowadays they constitute a common practice in most developed countries contributing to objective policy making. Although the agreement on the metrics to be used still constitutes a major debate to be resolved, both in the scientific and political spheres, consensus has been reached on the fact that the environmental performance of a product or process should be evaluated over its entire life cycle (Grossmann and Guillén-Gosálbez, 2010).

1.4.1 Lifecycle assessment (LCA)

LCA is a framework for identifying and evaluating the environmental burdens associated with a product or process over its entire life cycle (Guinée et al., 2002). Product environmental life cycle analysis (LCA) is used for identifying and measuring the impacts of different industrial products on the environment

If we have a production process like the one depicted in Figure 1.4, we will have as inputs the raw materials and utilities, and as outputs, the products, waste generated, and some emissions to air, soil and water. Before the raw materials and utilities enter the process, they might have gone through different pre-processing steps, such as extraction, storage and transportation, these are also tracked and their environmental impacts are accounted for in an LCA analysis. After the products exit the gate of the manufacturing plant, they might also experience different stages like the use and disposal/recycling phases. All of these steps have an impact in the environment, from what is called the cradle, to the gate or to the grave. The concept of cradle to grave analysis, which is a synonym of LCA, is known as well-to-wheels analysis when applied to transport fuels and vehicles. Thus, a well-to-tank analysis for transport fuels and vehicles is in effect a cradle to gate analysis.

In this work the cradle to gate analysis is adopted for the processing systems modelled in Chapter 3, Chapter 4 and Chapter 5, as indicated by the red rectangle in Figure 1.4. Subsection 4.1.1 gives an overview of the literature on environmental performance of process design and optimization and Subsubsection 4.3.6.2 introduces the LCA methodology.

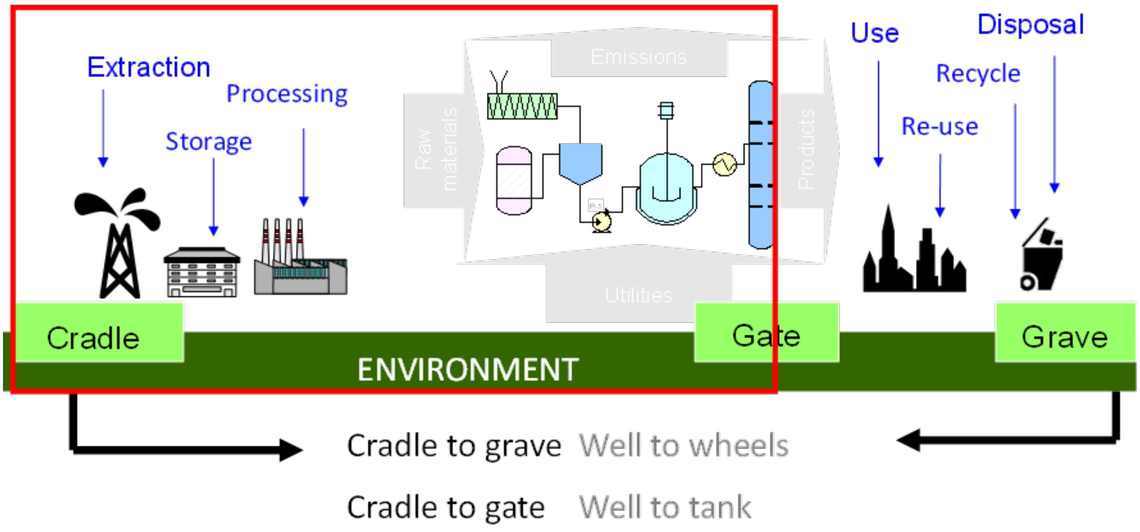


Figure 1.4: Scopes in life cycle assessment methodology

1.4.2 Environmental impact reduction: Multi-objective optimization (MOO) and LCA

Despite the advantages that LCA methodology brings for assessing the environmental impact of a product in a holistic manner, it lacks a systematic procedure to effectively compare multiple alternative routes in order inform and manage the reduction of environmental impacts. One possible manner to address this problem is by combining multi-objective optimization and LCA in a single mathematical programming framework (Azapagic and Clift, 1999). Figure 1.5 shows how the application of life cycle analysis to the optimization of a process or product is able to provide a holistic framework for reducing its associated environmental impacts. Traditional approaches focused on reducing the environmental impact associated to local stages of a product chemical supply chain, such as the production step. Therefore, by neglecting the interconnections of raw materials, production processes and uses of a product life cycle one may end reducing the impact locally in one stage at the expense of increasing the impacts in other stages of the product life. In contrast, the framework proposed here, by adopting a global boundary for the environmental impact quantification, avoids shifting the impacts to other stages of the manufacturing chain. In Figure 1.5 we can see how metric 2 is minimized locally at the manufacturing stage, while its impact in the rest of the stages continues to be high. Also Metric 1 is extremely reduced in the use and disposal phase, while its impact at the manufacturing stage escalates. On the other hand, by reducing the life cycle impact of a product, which is represented by metric 3 in the figure, a whole system reduction is achieved, which is represented in the right hand side bars representing the sum of the impacts along the different processing stages for each metric. This global approach avoids shifting the impact to other stages of the manufacturing chain, as opposed to traditional local environmental impact methods (see Figure 1.5).

The problem of simultaneously optimizing the economic and the life cycle environmental performance associated to a product or process system, gives rise to a multi-objective optimization problem whose formulation can be presented as follows.

$$\begin{aligned} (MOOe) \qquad & \min_{x,y,z} \left(C(x,y,z), EI_1(x,y,z)...EI_n(x,y,z)\right) \\ s.t. \qquad & h(x,y,z) = 0 \\ & g(x,y,z) \leq 0 \\ & x \in \Re,\ y \in \{0,1\},\ z \in N \end{aligned} \tag{2}$$

By solving $MOOe$, one can simultaneously optimize the economic and the environmental performance according to different environmental metrics. Some authors have adopted a similar approach for different supply chain design problems in the area of PSE (Hugo and Pistikopoulos, 2005; Bojarski et al., 2009; Guillén-Gosálbez et al., 2008, 2010), but despite these efforts there are open issues, like which indicators should be used, how can uncertainty be best managed and how to apply this approach to new problems. More specifically, in Chapter 4 and Chapter 5 a case where three environmental metrics addressing very different types of impacts are be optimized together with the economic performance of an industrial process system. On the other hand, Chapter 3 of this thesis presents a slightly different perspective of the problem, where eight different environmental impact metrics are simultaneously optimized together with the economic performance and then reduced to a reference subset of non-redundant representative metrics for the design of a hydrogen supply chain. Details of this last approach are presented in the following section.

The solution of MOO problems is given by an ensemble of Pareto points that form the Pareto set. Our goal is then to unveil the shape of the natural trade-off that lies behind the corresponding economic and environmental performance pairs and represent it in a two dimensional space. A review of the literature in the area of multi-objective optimization and LCA is given Subsection 4.1.2 and the details of the formulation of environmental impact categories within the objective functions of the problems are presented in Subsubsection 3.3.5.6 and Subsubsection 4.3.6.2.

1.4.3 Objective reduction: Multi-objective optimization (MOO) and Principal Component Analysis (PCA)

One of the major challenges in multi-objective optimization is the visualization and analysis of the multidimensional Pareto sets arising from the solution of problems with more than two objective functions (Branke et al., 2012). In these problems with many objectives, decision-makers need to deal with a large number of criteria, that might or might not be conflicting, but whose relationships at this stage are uncovered. When several objectives are non-conflicting, they can be considered are redundant since they

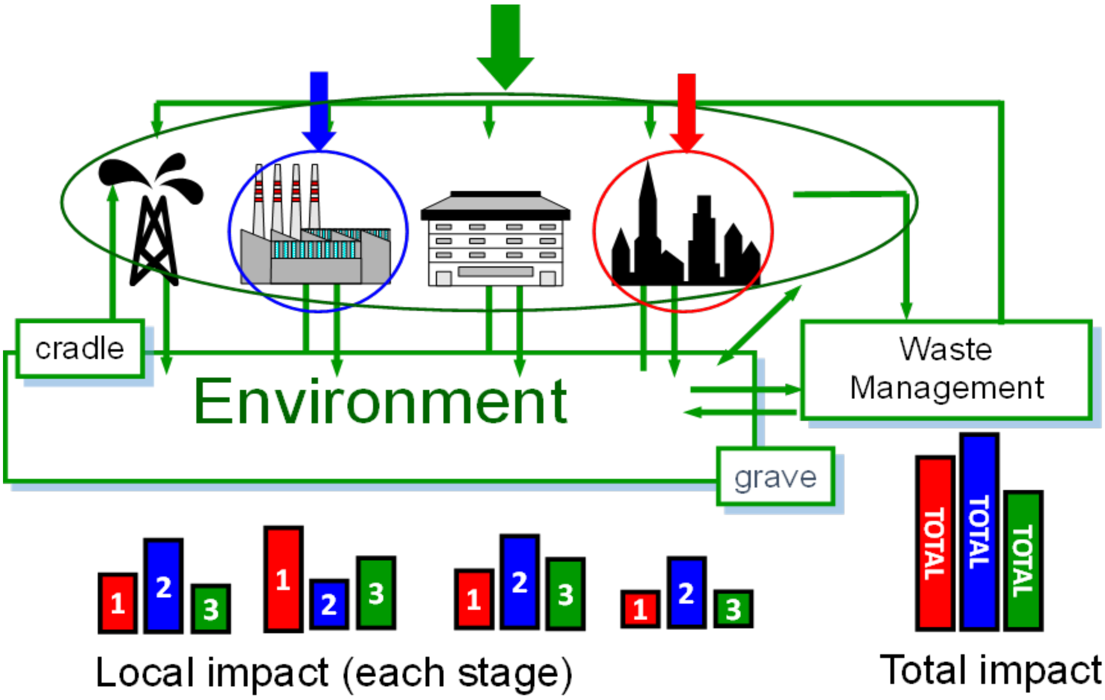


Figure 1.5: Holistic framework combining multi-objective optimization (MOO) and LCA. In red the environmental impact measured by metric 1 for each manufacturing stage: raw material extraction, manufacture, storage, distribution and use; in blue, the environmental impact according to metric 2 is shown. In green, environmental impact for metric 3. The total values represent the sum of each environmental impact metric for all the product life cycle stages are shown.

would effectively produce the same design solutions. Therefore the best practice in this case would be to remove them from the analysis.

Despite the advances in the MOO field on dimensionality reduction (Deb and Saxena, 2005; Brockhoff and Zitzler, 2009; Branke et al., 2012), its application to environmental problems is quite scarce in the literature. In a seminal paper Guillén-Gosálbez (2011a) presented a method for objective reduction based on a rigorous MILP formulation to systematically identify redundant objectives. Later, this approach was applied to different problems. For instance, Oliva et al. (2013) superposed a clustering algorithm to the original MILP and applied it to the design of petrochemical, hydrogen and bioethanol supply chains. Later,Vaskan et al. (2014) applied the MILP reduction technique to produce more sustainable planning and operation strategies of electricity and steam utility plants, and Copado et al. (2014) presented an algorithm for its application to environmental and systems biology problems. The advantage of this approach is that it can ensure a minimum distance error for the reduced dimensional set obtained, but its computational implementation can become challenging.

More recently, Kostin et al. (2015) implemented the MILP rigorous algorithm within the MOO resolution strategy that overcomes the need for post-optimal solution analysis. Antipova et al. (2015) used a post-optimal “Pareto filter” technique for the environmental MOO problem of a reverse osmosis plant. Despite being computationally efficient, this technique could not guarantee the best redundant objectives are selected.

In this work, we use a statistical technique known as Principal Component Analysis (PCA) for reducing the dimensionality of the objective set in a multi-objective optimization problem involving nine different objectives. PCA is a mathematical procedure that uses orthogonal transformation to convert a set of possibly related variables into a set of uncorrelated variables called principal components.

This orthogonal transformation is done by first standardizing the data (*i.e.*, subtracting the mean and dividing by the standard deviation) in order to find the correlation matrix and its eigenvectors (*i.e.*, e1, e2, e3). Next, the eigenvectors are ordered and these are called the principal components. The original dataset is finally projected in a smaller dimension defined by the eigenvectors (see Figure 1.6) and the principal components are found on that new basis. More details on the theory and the formulation implementation are found in Chapter 3. More specifically, a review of the literature on multi-objective optimization and PCA is introduced in Subsection 3.1.2 and the details of the methodology and its implementation in the mathematical formulation are given in Section 3.4. The results of its application to the design of a hydrogen supply chain in Spain are finally presented in Section 3.5.

Deb and Saxena (2005) investigated the use of PCA to identify redundant criteria in MOO problems. Then, Sabio et al. (2012) and Pozo et al. (2012) applied this approach for reducing the environmental objectives hydrogen infrastructure design and in petrochemical supply chains respectively. This framework proved less computationally intensive that the MILP rigourous approach, while still being able to ensure that the

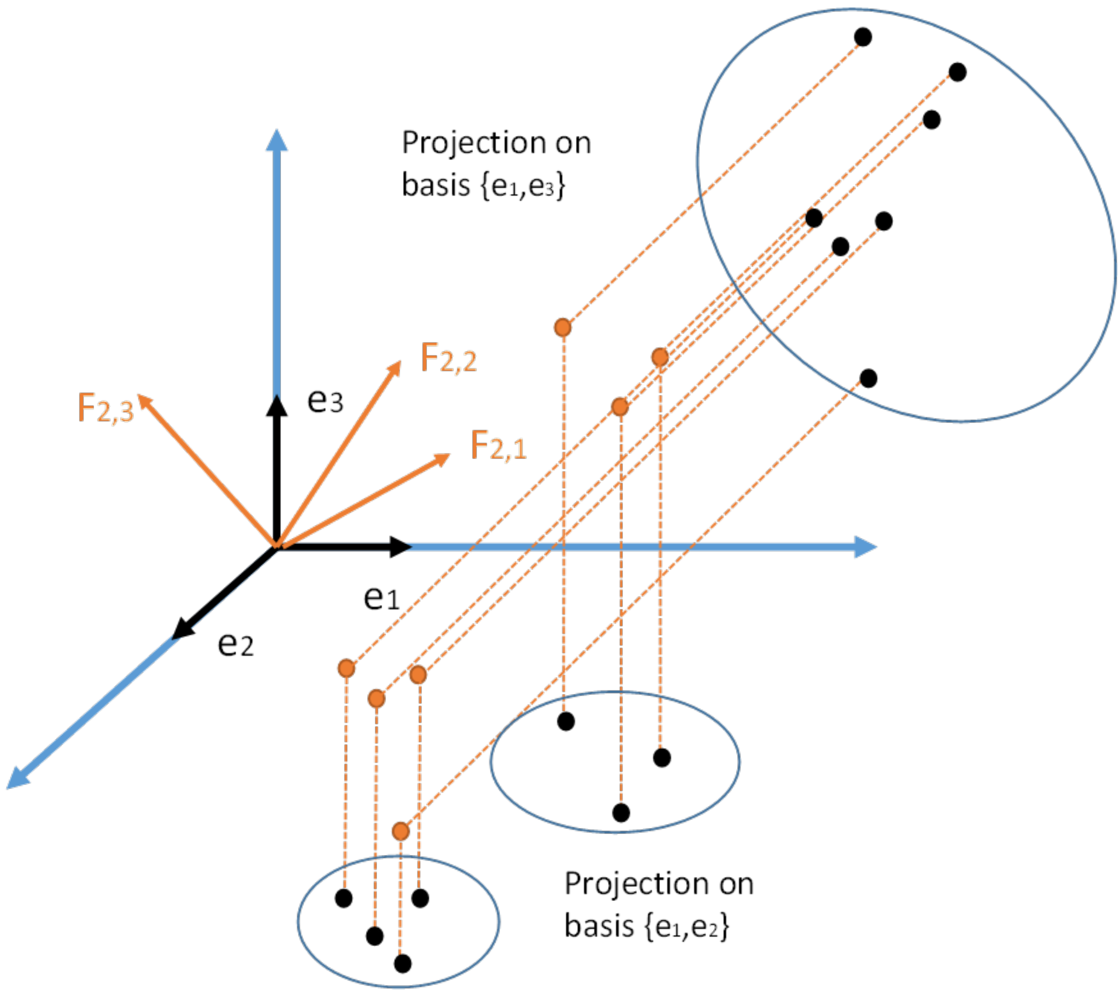


Figure 1.6: Orthogonal projections in principal component analysis

best objective set is selected according to the the statistic criteria used within the PCA setting.

1.5 Uncertainty

Optimization under uncertainty is the area that explicitly deals with the problem of incorporating the effects of "non-deterministic" information into the PSE mathematical programming frameworks. Its goal is to reflect real situations when decisions need to be made, even when their consequences are not fully understood.

Traditionally, the methods developed in formal sciences, such as mathematics, find their way around uncertainty by providing accurate problem definitions (*i.e.*, hypothesis, axioms) that restrict the problem environment to strict pre-defined conditions. This is the main reason why problem definition is a key step in the scientific method. Nevertheless, when formal sciences are applied to real problems, such as the ones found in biology, engineering or chemistry, they confront the need to deal with data and experiments, which at the same time depend on the states or conditions under which those data are obtained. At this stage it is easy to observe how the effects of uncertainty start to filter through the elements of the system under study to manifest as parameter variations in this particular case. Not in vain, the original word for 'statistics' was coined in part to define a field that dealt with information about 'states'.

On the other hand, the areas more closely related to systems sciences, which in our case are represented by management science and life cycle analysis, deal with environments that are less easily reducible to a controlled unit (*i.e.*, humans, environment). Although in these cases variations in data are also observed, their causes cannot be easily unveiled by just changing the states or conditions of the experiments. For instance, methodological and model uncertainty turns out to be of particular importance in these fields, since even the frameworks used to define the concepts are highly debated. Therefore, parameter uncertainty, although still existent, is overshadowed by the magnitude of uncertainties affecting larger system scales. An overview of uncertainty characterization and quantification in LCA is presented in Subsection 5.1.2 and Subsection 5.1.3.

In contrast, in formal and natural sciences, where the theories are perhaps more easily testable and therefore accepted, the theoretical frameworks can be taken as "constant" and therefore the major part of the uncertainty analysis is concentrated on parameters. Parameter uncertainty is the uncertainty associated to the parameter values, usually taken from economic, demographic, epidemiological or experimental studies (Bilcke et al., 2011). To overcome the limitations associated with this perspective, sensitivity analysis and model testing are stressed out as activities of high relevance and decisions are frequently taken by experts in the systems. Despite these differences, science-based and systems-based research fields have in common the necessity to develop models and frameworks able to represent the realities of the problems they study, which take a central role in both fields. This is portrayed in the echelons represented in Figure 1.1.

Given the pervasive nature of uncertainty, it is impossible to imagine how models themselves would escape from it. Structural or model uncertainty is related to the aspects of the problem (*i.e.*, constraints) represented in the model and the model paradigm (*i.e.*, simulation or optimization, deterministic or stochastic). Examples of structural uncertainty within optimization frameworks are choice of available technologies in a flowsheet design task, the use of capacity expansions in supply chain problems or the use of single period or multi-period problem formulations. Thus, the approaches to deal with structural uncertainty can range from model reformulations (*i.e.*, addition and removal of constraints and functions) to the comparison of different modelling frameworks for a similar case study.

Methodological uncertainty in contrast arises from differences used in the evaluation methodology (Bilcke et al., 2011). Dealing with uncertainty due to choices can involve then evaluating different perspectives taken in a decision-making activity (*i.e.*, consumer

Table 1.1: Examples of uncertainty in each phase of process systems design in PSE

Phase / Source	Data collection	Superstructure Postulation	Model formulation	Model solution
Parameter uncertainty	Demand Energy prices Technology costs Technology efficiency Correlation value			
Model uncertainty		Design problemas: Available process units/technologies Potential process connections Capacity expansión (yes/no) System echelons	First principles/ Experimental correlations Single/Multi-period, Deterministic/Stochastic Probability distribution type Correlation (yes/no)	
Methodological uncertainty			Objective function/perspective: Producer/consumer, Economic/environmental Risk averse/Opportunistic	Weights Allocation method Time horizon

vs producer, environmental vs economic), the valuation technique (*i.e.*, single-objective or MOO, weights), the time horizon for the analysis or allocation methodologies in LCA. Examples of all these types of uncertainties found in PSE optimization problems are presented in Figure 1.1.

In sum, we have that science-based research has traditionally focused on parameter (or parametrizable) uncertainty (Sahinidis, 2004), while systems-based research has dealt more deeply with the issues of defining and classifying uncertainty (Huijbregts, 2001; Bilcke et al., 2011). While the aim of this thesis is not to provide a philosophical framework for understanding uncertainty, the systems nature of the processes under study provides a good case for applying the knowledge gained from the study of uncertainty in both, science-based and systems-based research. Arguably, we consider that optimization as a mathematical framework already deals with an important part of the model structural uncertainty by including all possible alternatives and choices available for each problem.

More specifically, this work analyses the uncertainty related to energy prices in Chapter 2 by implementing a stochastic probabilistic scenario framework combined with risk management for controlling the variability of the economic performance at the design stage of a hydrogen supply chain. This proves to give robust solutions able to cope with this type of uncertainty. In Chapter 4, we formulate a novel MOO deterministic MINLP framework for the design of process industries, where three environmental metrics are optimized along with the economic performance. In this case, two different energy price data sets were tested, fixed and flexible demand satisfaction and extreme tests and sensitivity analysis were performed to check the limiting constraints of the problem. In Chapter 5, the uncertainty associated to the life cycle inventory of emissions of the industrial plant is considered using scenarios that can model any type of continuous/discrete probability distributions. Particularly, without loss of generality, we focus on lognormal distributions that are able to capture the high uncertain, but less probable scenarios by means of their associated long tails. All these analysis can be mapped on to different uncertainty analysis categories, as it is explained in the following sections.

1.5.1 Parametric uncertainty: sensitivity analysis and stochastic programming

Parameter uncertainty, as introduced by Bilcke et al. (2011) appears when there is uncertainty in the choice of parameter values. Traditionally, in mathematical programming the frameworks considering that all parameters are known in advance are said to be deterministic. This type of problems are characterized by a single solution point that describes the optimal design or plan that the decision-maker should adopt (see Figure 1.7a). In contrast, when uncertainty is considered, the solution to the problem might be given a range of solutions corresponding to different process system designs in our case.

Optimization under uncertainty is the field that deals with the development of systematic mathematical programming frameworks capable of considering uncertainty at the modelling stage. The two major mathematical programming approaches used for process design under uncertainty are robust optimization and stochastic programming. Robust optimization is more widely applied on short term scheduling problems where ensuring the feasibility of the constraints over a given uncertainty range is of primal importance. On the other hand, stochastic programming can better handle flexibility *via* the recourse actions taken once the uncertainty is revealed in a large number of scenarios, being therefore better for longer term strategic problems (Grossmann, 2014).

It is worthwhile to mention that stochastic problems present major computational challenges associated to the existence of scenarios and scenario trees. Robust optimization and chance constraint formulations, on the other hand, do not suffer from these problems, but still present challenges for the computation of the probabilities and their associated derivatives, and lack in turn the ability to represent recourse actions. In this thesis, both the supply chain design and the industrial plant synthesis are long-term problems of at least one year planning horizons, and in consequence a stochastic programming formulation has been considered as better suited for addressing their parametric uncertainties.

Both of the approaches introduced in the previous lines give rise to sophisticated systematic tools for addressing uncertainty in a rigorous and robust manner, as they either involve the integration of full probability distribution functions or the sampling of as many scenarios as it is required to ensure the solutions lie within an acceptable confidence level. More classical methods for dealing with parameter uncertainty either rely on manually changing few parameters, most commonly known as sensitivity analysis. In this work we consider that parameter uncertainty approaches are those which are mainly concerned with the uncertainty related to parameter values, as it is the case of stochastic programming, robust optimization or sensitivity analysis. Examples of parameter uncertainty found at different stages of process systems design in PSE are given in Figure 1.1.

Although stochastic programming and robust optimization approaches involve considerable problem reformulations affecting the model structure, as their focus here is to assess the effect of parameter variations in the solution, they are classified as techniques for dealing with parameter uncertainty. In this work, the main approaches used for parameter uncertainty are manual parameter change, sensitivity analysis, complete dataset change and stochastic programming (see Figure 1.2). In particular, uncertainty in energy prices was addressed *via* a stochastic programming formulation presented for the case of a hydrogen supply chains design (see Subsection 2.3.1). An overview of the literature on hydrogen infrastructure optimization under uncertainty can be found in Subsection 2.1.3. In Chapter 4 the different parameter tests are performed for the new deterministic industrial synthesis design problem presented. More specifically, the model is run for different energy datasets in Subsection 4.5.1 and Subsection 4.5.2 and an extreme test is run to determine the limiting product quality constraints in order to run a sensitivity analysis on them (see Subsubsection 4.5.1.4). In Chapter 5 the uncertainty in the life cycle inventory of emissions for the industrial plant design problem is presented through a stochastic formulation in Section 5.3 and the results also explore the effect of different combinations of correlated uncertain parameters, as introduced in Subsection 5.5.1. An overview of uncertainty in LCA and optimization is introduced in Subsection 5.1.1 and of multi-objective optimization under uncertainty in Subsection 5.1.4.

1.5.2 Structural uncertainty: constraint reformulations and deterministic-stochastic analysis

Structural uncertainty relates to the type of model and structure chosen for the analysis (Bilcke et al., 2011). In this step, the expert needs to decide which parameters, variables and equations are going to be represented in the problem formulation. As an example, in Chapter 2 and Chapter 3 the hydrogen supply chain design problem is represented, in very general terms, as a mass balance constrained problem for the specific geographical locations and a capacity expansion coupled with production rates for the manufacturing stages. Accordingly, in Chapter 4 and Chapter 5, the representation of the manufacturing plant is done using the concepts of unit operations, experimental correlations, flow diagrams and process synthesis(McCabe and Smith, 2001; Sargent, 2005; Grossmann, 1989).

Structural uncertainty analysis is therefore not focused on the parameter variation effects, but on the effect of the model structure in the results obtained. For instance, it is well known in PSE that the formulation of first principles models poses significant challenges, in particular for large scale process systems operations like chemical complexes or even complex process units (*i.e.*, distillation columns). To overcome these problems,

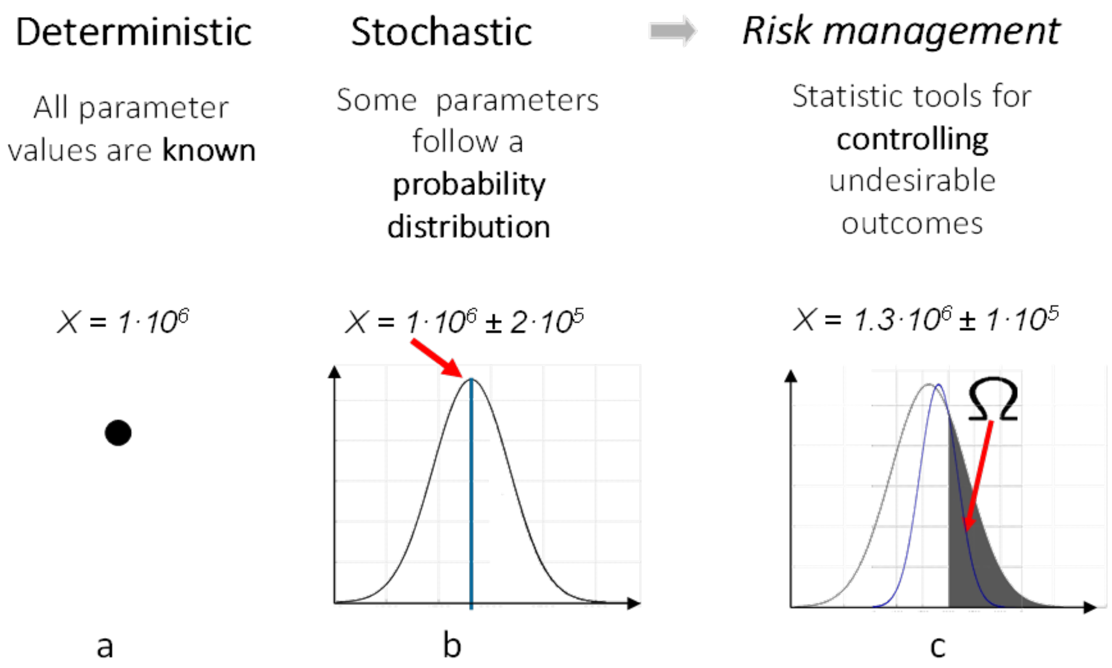


Figure 1.7: Solution representation for deterministic, stochastic and risk management frameworks. (a) Single point solution for a deterministic problem. (b) Expected value and interval range of the probability distribution output of a stochastic probabilistic framework. (c) Expected value and interval range of the probability distribution output of a stochastic framework in including risk management for reducing the probability of obtaining outcomes higher than Ω.

simplified short-cut models or process linearization are often pursued as the most popular strategies to resolve the models. Experimental correlations can be seen also as a simplified representation of the processes, although given their uncontrollable nature, they might present mathematical forms that give rise to similarly challenging problems.

Other common discussion in the area of PSE is the single-period (*i.e.*, static) or multi-period (*i.e.*, dynamic) representation of process models. Despite the advantages of formulating multi-period problems, these tend to increase the computational challenges of their single-period counterpart by several scales of magnitude particularly in large-scale complex models. In a similar line, the issue of deterministic and stochastic models poses significant implementation challenges, but these have been indeed raised in order of importance given the impacts, current relevance and pervasive nature that uncertainties have in PSE (Grossmann, 2014). An attempt to map these uncertainties in the different process system design stages is presented in Figure 1.1.

Note that here we refer only to structural uncertainty within the optimization field, but structural uncertainty analysis should also be addressed by comparing different modelling frameworks (*i.e.*, simulation vs optimization), although the modelling paradigms by themselves pose significant challenges to these tasks. For instance, whereas in a simulation framework a structural uncertainty analysis might involve the consideration of a different process unit to perform a task, in optimization frameworks theoretically all possible process units are included. In this work, structural uncertainty is addressed very modestly in three different forms.

- Allowing flexible demand (*i.e.*, changing an equality by an inequality constraint).
- Including and excluding uncertain parameter correlation.
- Solving an MOO model for economic and environmental performance measured *via* Metric 1 in deterministic and stochastic form.

Here, we mention a stochastic problem formulation as an approach for analysing structural uncertainty. This is due to two main reasons: first, the focus here is on detecting differences that the structure of the model formulation produces in the results, and second, the same model is solved in a deterministic and stochastic formulation sequentially. The results of these analysis can be found in Chapter 4, Subsection 4.5.1, and Chapter 5, Subsection 5.5.1 and their respective Pareto sets are presented in Figure 4.8 and Figure 5.5. The results of reformulating flexible and fixed demand satisfaction constraints had a larger impact than expected and are shown in Subsubsection 4.5.1.3, whereas the uncertain environmental parameter correlation effects fall in a different impact scales depending on the metrics used, as shown in Subsection 5.5.1, Figure 5.5 and Figure 5.6. Note that these methods prove to be a useful practice to explore the deterministic model boundary conditions, but do not provide a way of systematically generating robust solutions. The latter is only accomplished by the use of the stochastic programming formulation presented in the previous section, where the uncertainty can be embedded at the design step.

1.5.3 Methodological uncertainty: multi-objective optimization and Pareto sets

Uncertainty around methodological choices arises when different views exist about what constitutes the correct approach for optimum decision-making (Bilcke et al., 2011). Examples for economic evaluation include the perspective taken (*i.e.*, risk averse or opportunistic decision-maker, economic performance, environmental impact, societal welfare), whereas in life cycle analysis these take usually the form of impact allocation procedures, weights in integrated impact indicators or the specific environmental impact metric chosen (*i.e.*, carbon, carbon equivalent, greenhouse gases, acidification and eutrophication, respiratory effects on humans). Huijbregts (2001) recognised this category as uncertainty due to choices in his characterization of uncertainty in LCA. Examples of these uncertainty type in PSE are presented in Figure 1.1.

In the present work, we exploit the capabilities of stochastic programming with recourse and enlarge its scope by introducing the concept of risk management represented by the worst case and downside risk metrics in the objective function (Barbaro and Bagajewicz, 2004; Bonfill et al., 2004; Eppen et al., 1989). These risk metrics, traditionally more used in robust optimization approaches, are applied here to two long-range planning design problems in Chapter 2 and Chapter 5. Subsubsection 2.3.6.6 gives an overview of the risk management framework used and Figure 5.2b portrays the differences in both risk metrics. The aforementioned worst case formulation allows the representation of allows the representation of risk-averse decision-makers in stochastic programming, but it can be considered an overly conservative metric in some instances (Grossmann, 2014). For this reason, the downside risk as defined by Eppen et al. (1989) is appended to the stochastic industrial problem formulation in Subsection 5.3.1 and considered as an additional risk metric to be optimized together with the traditional expected performance and the worst case metric.

In contrast to their deterministic counterpart, stochastic problems present a range of solutions that in their simplest form are reduced to the corresponding expected value. The effect of substituting the expected value by a risk metric is represented in Figure 1.7b and Figure 1.7c, where its ability to control the undesired outcomes is portrayed. The resulting multi-objective formulation of these problems has the form of $MOOr$ in Eq. 3, where the first objective function is the expected value of the system cost and the second is the worst case.

$$
\begin{aligned}
(MOOr) \qquad & \min_{x,y,z} \left(EC(x,y,z), RM_1(x,y,z)...RM_n(x,y,z),\right) \\
s.t. \qquad & h(x,y,z) = 0 \\
& g(x,y,z) \leq 0 \\
& x \in \Re,\ y \in \{0,1\},\ z \in N
\end{aligned}
\tag{3}
$$

As mentioned before, Chapter 2 and Chapter 5 of this thesis use this combined approach applied to a hydrogen supply chain design and planning problem and to the synthesis and operation of an industrial process plant respectively. In Subsubsection 2.3.6.1 and Subsubsection 2.3.6.6 the implementation of the *MOOr* multi-objective formulation, where the expected economic performance and the worst case value are simultaneously optimized, is presented for the MILP problem on hydrogen supply chains design in Spain. In Subsection 5.3.1 the stochastic formulations for the expected economic performance, risk management, worst case value and downside risk metrics are depicted for the multi-objective MINLP synthesis problem of an industrial chemical plant. Next in Subsection 5.3.2, Subsection 5.3.3 and Subsection 5.3.4 the scenario planning probabilistic "quasi-simulation" approach used to for the stochastic problem resolution is introduced in detail. The problem solutions are presented first *via* their respective Pareto sets in Figure 2.3, Figure 2.4, Figure 5.5 and Figure 5.6, and the solution visualization is complemented by the use of the cumulative probability curves for each objective in Figure 2.7, Figure 5.7, Figure 5.8 and Figure 5.9. The latter set of figures provide a picture where all the objective function profiles can be combined, thereby allowing to compare different risk values associated to the design solutions obtained when using the expected performance, worst case value and downside risk metrics respectively. The types of uncertainty addressed in each part of this thesis are presented in Figure 1.2.

The selection of the metric to be optimized is a key element in optimization. Traditional performance indicators in chemical process design have focused on economic metrics, like profit or costs, and on non-economic objectives, such as customer satisfaction or product quality. However, organizations have recently started to focus more on improving their environmental performance. One could argue that costs for environmental considerations could be derived and therefore implemented as additional constraints in a single-objective framework. However, assessing these costs (externalities caused by the chemical process) is very challenging. Hence, MOO is better suited to deal with this problem, as in addition to avoiding the need to calculate externalities it allows identifying solutions where significant environmental improvements can be attained at a marginal increase in cost.

Pareto sets are an effective representation for multi-objective problems, for they allow to graphically visualize the trade-off of conflicting objectives. In fact, by looking at Figure 1.8, we can see that after both metrics have been optimized independently, the Pareto front can inform of the costs associated (*i.e.*, reduce F_1) for reducing a given amount the environmental impact (*i.e.*, reduce F_2). The figure depicts how the Pareto frontier of an integer linear or non-linear problem looks like, where the points represent the problem design solutions, and the lines are just artificially added to connect the discontinuous Pareto set obtained when solving a MILP or MINLP problem. The extreme point depicted in light green represents the solution that independently minimizes F2, while the black point on the other end represents the solution that minimizes F1. All the green points in between these extreme solutions, and in the case of our synthesis problems, reflect the optimal non-dominated (*i.e.*, Pareto efficient) feasible designs that

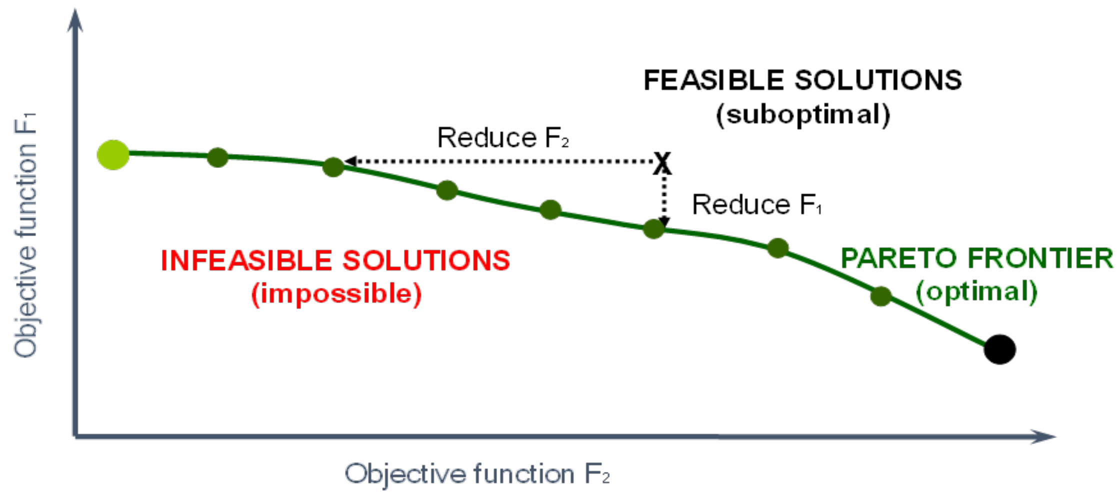


Figure 1.8: Pareto set as a tool for uncovering methodological uncertainty in optimization frameworks

simultaneously optimize F_1 and F_2. The advantage of this methodology is that it ensures that the decision maker selects one option located in the frontier of efficient or non-dominated solutions by default. That is, to guarantee that the decision is optimal in the resulting two dimensional space. This can be generalized for the case of more objectives and other objectives than cost and environmental performance.

We argue here that multi-objective optimization is an effective methodology for dealing with the uncertainty associated to the methodological choice of performance indicators in mathematical programming. This is because it allows assessing the impact of the choice of a particular metric on the final Pareto solution selected. Not in vain, Grossmann (2012, 2014) recently highlighted the importance of research on optimization under uncertainty in PSE. Traditional techniques used in optimization under uncertainty are also effective for addressing structural uncertainty, but that needs to be considered more explicitly at the modelling stage. The main goal of the uncertainty framework outlined here is to help PSE practitioners to implement systematic procedures for dealing with whole system uncertainties at the modelling stage (see Figure 1.1), thereby achieving more robust process models able to cope with all types of systems uncertainties.

Table 1.2: Uncertainty types and approaches used

Uncertainty type	Thesis chapter	Example/application	Approach
Parameter uncertainty	Chapter 2	Energy prices	Stochastic programming – Gaussian probability function
	Chapter 3	X	X
	Chapter 4	Energy data set	Manual dataset change and model run
		Limiting product quality constraints	Extreme test and sensitivity analysis
	Chapter 5	Lifecycle inventory of emissions	Stochastic programming – Lognormal probability function
		Correlated parameters	Different correlation combinations (i.e., Case 2, Case 3)
Model uncertainty	Chapter 2	X	X
	Chapter 3	X	X
	Chapter 4	Demand satisfaction: flexible/fixed	Manual constraint reformulation and model run
	Chapter 5	Correlation (yes/no)	Manual reformulation of two cases (i.e., Case 1: no correlation, Cases 2&3: high correlation)
	Chapter 4&5	Environmental impact uncertainty	Deterministic vs stochastic model runs for Metric 1
Methodological uncertainty	Chapter 2	Economic risk perspective	MOO: Expected cost versus economic worst case
	Chapter 3	Environmental performance metric	MOO: Cost versus 4 damage categories and 8 environmental impact metrics using Eco-Indicator 99 methodology
	Chapter 4	Environmental performance metric	MOO: Economic performance versus 3 different environmental performance metrics.
	Chapter 5	Environmental risk perspective	MOO: Cost versus Expected environmental performance, downside risk and worst-case metric

1.6 Outline

The capabilities of the systematic multi-objective optimization frameworks arising from the combination and individual use of these techniques is demonstrated through its application to two major design problems.

The first half of this thesis is devoted to the design and planning of hydrogen supply chains for vehicle use, while in the second half we introduce a novel formulation for the synthesis and operation of a complex industrial process system. The modelling tasks are carried out by initially ensuring the efficient solution of the deterministic formulation and next proceeding to the formulation and solution of their specific stochastic counterparts stochastic as prescribed by Grossmann (2014). Note however, that the order followed to present the chapters has been chosen by chronology of their publication order, which does not necessarily reflect the order of the modelling tasks carried out. Thus, if we look at Figure 1.1, we therefore navigate from the larger geographical and temporal scales involved the in multi-site design problems, to the smaller temporal and geographical scales involved in the design and operation of a single-site problem, following the "uncertainty and abstraction" arrow, and in increasing order of formulation complexity indicated by the "precision" arrow.

Hydrogen supply chain

In Chapter 2 we introduce the first MOO stochastic framework to effectively deal with the problem of controlling the uncertainty of the economic performance arising from volatile energy prices on the design and planning of hydrogen supply chains for vehicle use in Spain (see Section 2.2 and Section 2.5). This problem is of particular relevance for the energy challenge mentioned at the beginning of this chapter, since hydrogen still continues to be an alternative to fossil fuels showing high flexibility for addressing the transition to a low carbon economy, but also with the potential of being increasingly turned into a renewable and zero-carbon energy vector (see Subsection 2.1.1, Subsection 2.1.2 and Subsection 2.1.3). In particular the model includes the formulation for all possible road and maritime transportation modes, including ships, trains, trucks and hydrogen pipelines. The production routes modelled include steam methane reforming with carbon capture, coal gasification with carbon capture, biomass gasification and electrolysis. The storage options include cryogenic spherical tanks for liquid hydrogen and pressurized cylindrical vessels for compressed hydrogen. The formulation of the resulting three echelon supply chain gives rise to a multi-objective stochastic multi-period problem that minimizes the expected cost and the worst case metric to address the financial risk of the system. The corresponding mathematical formulation is presented with detail in Section 2.3. A two-step algorithm is also devised and introduced in Section 2.4 for the resolution of the problem, which reduces its associated computational time in one order of magnitude. For more information see the corresponding scientific publication by Sabio et al. (2010).

Chapter 3 returns to the deterministic formulation of a hydrogen supply design problem (see Subsection 3.1.1, Section 3.2 and Section 3.5), but this time including the life cycle inventory of emissions associated to eight individual environmental metrics that form part of the Eco-Indicator 99 methodology, and the total discounted cost of the system. The results show that the problem solutions follow similar patterns for several indicators. The arising visualization and analysis difficulties associated with the high dimensionality of the problem are addressed by identifying redundant life cycle environmental metrics using principal component analysis (PCA). As a result of the combination of MOO and PCA (see Subsection 3.1.2 and Section 3.3), the dimensionality of the problem is halved and four environmental metrics are identified as redundant. These results indicate that half of the initial set of environmental metrics are able to represent the eight initial criteria in the supply chain design space. More details can be found in the corresponding scientific publication by Sabio et al. (2012).

Industrial process plant

Chapter 4 presents a novel formulation for more sustainable synthesis and operation of a complex industrial plant (see Section 4.2). Here, LCA is brought together with MOO for the synthesis and operation of the industrial scale problem presented in Section 4.1. The operations intervening in the process units are defined by non-linear equations expressed through complex process-specific experimental correlations[2]. More details on the mathematical formulation are given in Section 4.3. The model allows for the use of external intermediate products and raw materials, changes in the internal flows of the network in order to by-pass specific process units (thereby modifying the current topology) and fuel switching as design options to meet the desired product quality constraints while optimizing the objectives considered. The design task is thus posed as a multi-objective formulation where the economic performance and three different environmental metrics are considered, and the corresponding Pareto sets obtained after the model is resolved are depicted in Section 4.5. In addition, Subsection 4.5.1, Subsection 4.5.2, Subsubsection 4.5.1.3 and Subsubsection 4.5.1.4 show the effects of using different energy price datasets, allowing flexible demand and performing sensitivity analysis around the most influential product quality constraints. The methodological uncertainty associated to the use of single-objective optimization frameworks is also addressed by first solving a bi-criteria problem for one environmental impact indicator and then adding other two environmental metrics to the analysis (see Subsubsection 4.5.1.1, Subsubsection 4.5.1.2, Subsubsection 4.5.2.1 and Subsubsection 4.5.2.2).

Finally, in Chapter 5 the deterministic formulation introduced in Chapter 4 is taken as a basis to develop a stochastic programming model. More specifically, addressing the uncertainties associated to the life cycle inventory of emissions gives rise to a multi-objective stochastic problem where the LCA uncertainty is addressed at the design stage. An overview of the literature in the fields of uncertainty in LCA and optimization under uncertainty is introduced in Subsection 5.1.1, Subsection 5.1.2, Subsection 5.1.3 and

[2]Note that the complexity of modelling a problem of that scale using first-principles makes is one of the major challenges that industries and consultancies are facing nowadays, and therefore is out of the scope of our research time frames and desktop computational capability

Subsection 5.1.4. Next, the industrial scale problem to be solved is defined in Section 5.2. The problem is formulated as a multi-objective framework where the economic performance is simultaneously optimized with three different stochastic and risk management metrics: the expected value of the environmental impact, its worst case value and the downside risk metric. The goal of this approach is to represent different decision-making perspectives towards risk. By using three different stochastic and risk management metrics, the model deals with the uncertainty associated with single-objective optimization problems (see Section 5.3). Additionally, three cases of increasing order of magnitude in the associated uncertainty levels in the life cycle inventory of emissions are explored. The model is solved using a multi-objective stochastic optimization approach introduced in Section 5.4 and results are presented only for one environmental indicator in this case, three cases for correlation levels of the probability distributions and the three different stochastic and risk management metrics in Section 5.5.

This industrial scale problem addresses the challenge of formulating systematic frameworks for large scale problem resolutions. The underlying computational challenges arise mainly from the pervasive non-convexities associated to the form of the process correlations. To our knowledge it is the first study bringing to industry an MOO-LCA systematic framework capable of managing uncertainty. More details on the study can be found in its corresponding scientific publication in the *AIChE Journal* (see Sabio et al. (2014)).

Figure 1.9 shows a map of the contents presented the thesis, where the methodologies used and problem characteristics are mapped to the corresponding chapters. Later, Figure 1.10 provides a more classical outline of the thesis, where the contents of each chapter and general structure of the different thesis parts are depicted.

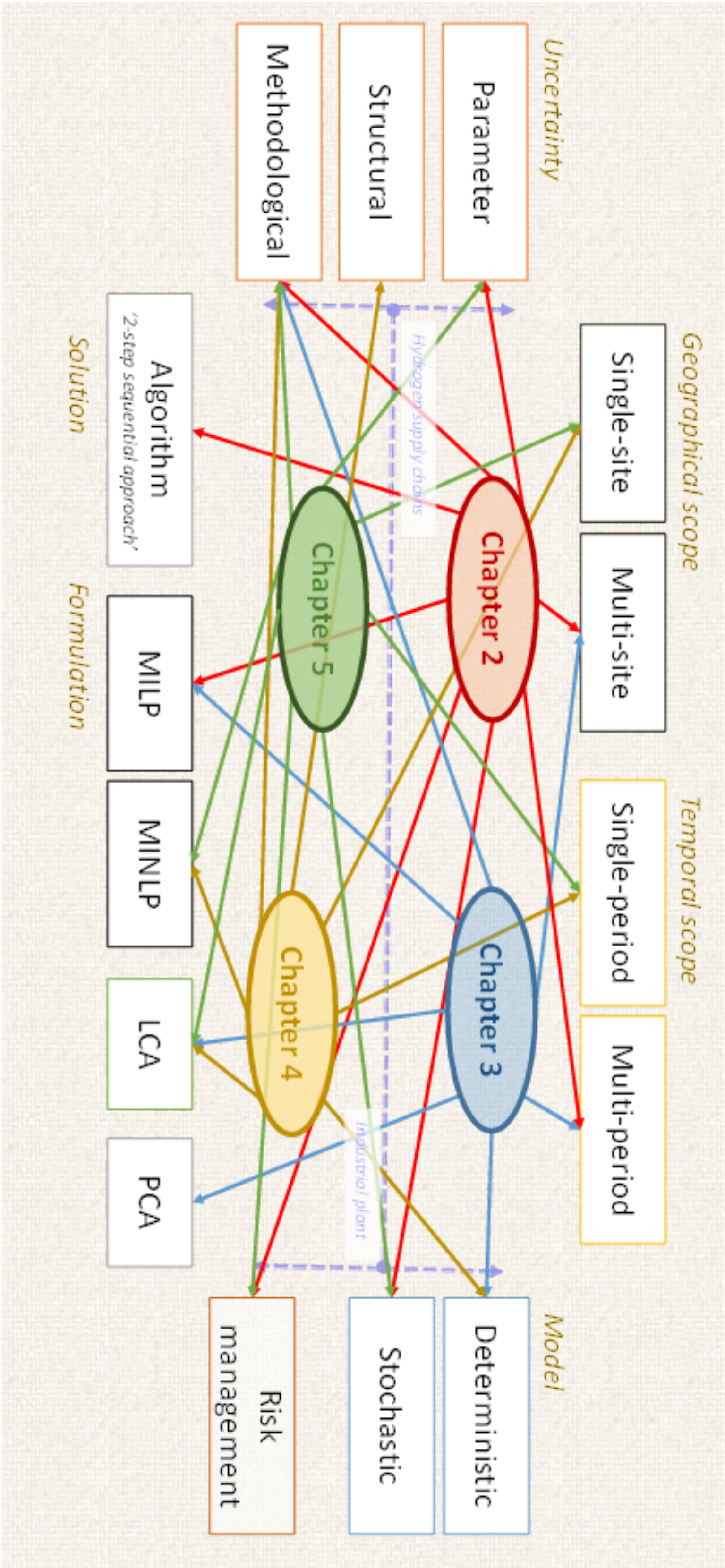


Figure 1.9: Thesis content map

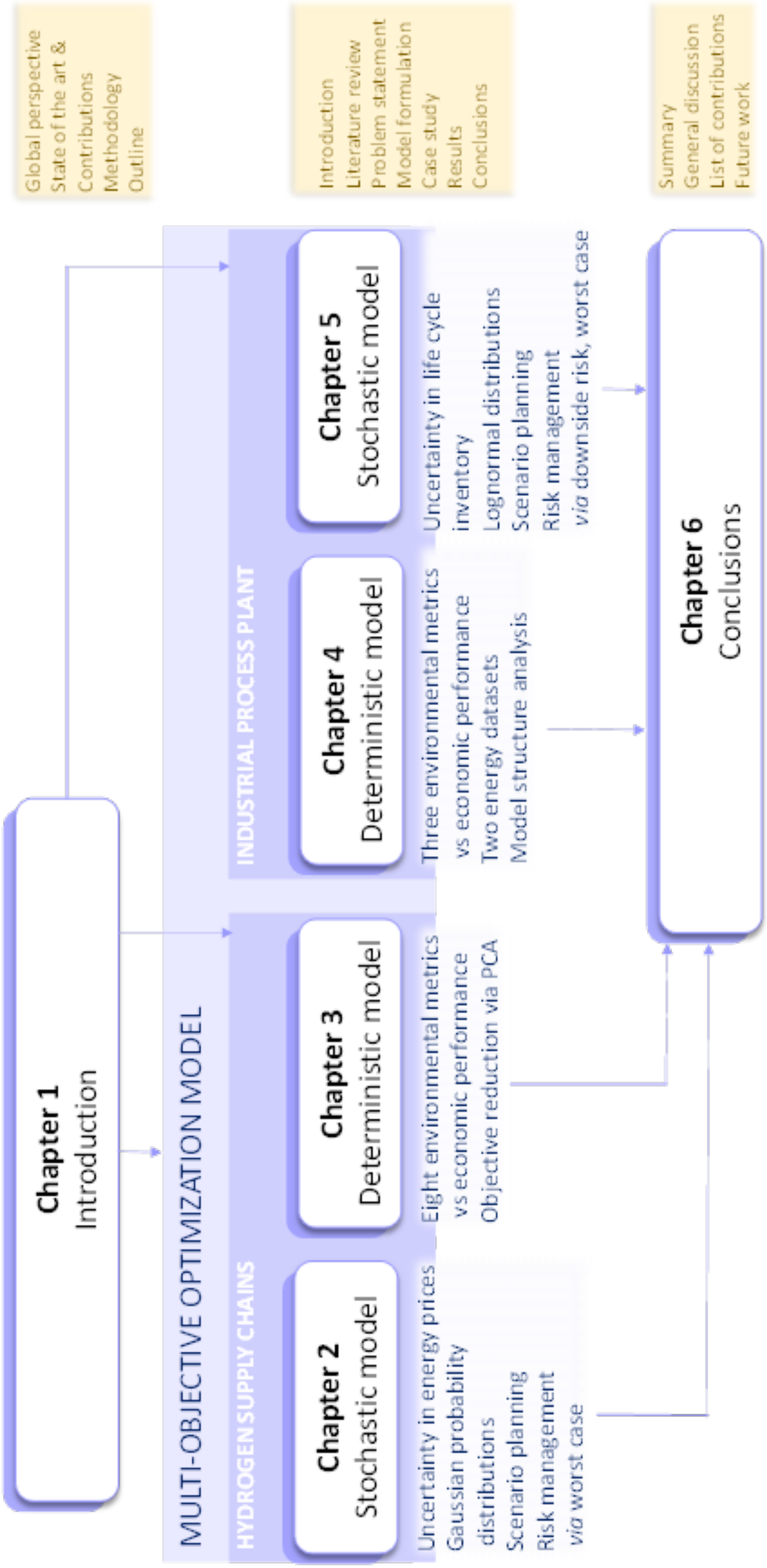


Figure 1.10: Thesis outline

Notation

Indices

e	scenarios
n	objectives

Sets

F	set of objectives
g	set of inequality constraints
h	set of equality constraints

Parameters

Ω	target for risk management

Variables

EI	Environmental impact
EC	Expected economic performance
C	Economic performance
RM	Risk metric
n	Number of environmental or risk metrics

CHAPTER 2

MULTI-OBJECTIVE OPTIMIZATION OF A HYDROGEN SUPPLY CHAIN – STOCHASTIC APPROACH

Water will one day be employed as fuel, that ***hydrogen*** *and oxygen of which it is constituted will be used*

Jules Verne, The mysterious island, 1874

2.1 Introduction

In this chapter we present a decision-support tool to address the strategic planning of hydrogen supply chains for vehicle use under uncertainty in the operating costs. This work is inspired on previous formulations of the hydrogen supply chain design problem (Guillén-Gosálbez et al., 2010; Guillén-Gosálbez and Grossmann, 2010) and extend their capabilities to represent different decision-making perspectives towards risk. Our aim here is to explore the capabilities of a stochastic optimization framework able to handle price volatilities in the energy sector for a higher level energy systems strategic planning and design task within the chemical supply chain. Given is a superstructure of alternatives that embeds a set of available technologies to produce, store and deliver hydrogen. The objective of our study is to determine the optimal design of the production-distribution network capable of fulfilling a predefined hydrogen demand. The design task is formulated as a multi-objective, multi-period and multi-scenario stochastic mixed-integer linear problem (MILP) that considers the uncertainty associated with the coefficients of the objective function of the model (*e.g.* operating costs, raw materials prices,*etc.*). The novelty of the approach presented is that it allows decision-makers to control for the effects of uncertainty in the economic performance at the hydrogen network design step. This is accomplished by using a risk metric that is appended to the objective function as an additional criterion to be optimized. An efficient decomposition method is also presented that proves to expedite the solution of the underlying multi-objective model in one order of magnitude by exploiting its specific structure. The capabilities of the proposed modeling framework and solution strategy are illustrated through the application to a real case study based on Spain, for which valuable insights are obtained. For more details on this work, please refer to the original publication (Sabio et al., 2010).

2.1.1 Prospects of hydrogen as a future energy vector

The growing concern about possible disruptions in the oil supply and the need to reduce greenhouse gas (GHG) emissions have fostered in recent years the research for a more sustainable energy and transport model ((Weissman et al., 2011, 2012),). Globally, the transportation sector accounts for an 18% of carbon dioxide global emissions and a 25% of the primary energy use (World Resources Institute, 2005; WRI, 2005). Within this context, hydrogen seems a potential alternative fuel and energy carrier since it can be produced safely and locally, besides having the possibility of being environmentally friendly. Not in vain, several books and information sources have been devoted to explain the opportunities and challenges of a future hydrogen economy (Ball and Wietschel, 2009a,b; Dunn, 2002; Ball et al., 2007; Kim and Moon, 2008a).

Recent views suggest that the transition to the hydrogen economy will depend on two main factors that must be developed in parallel:

1. Construction of an efficient hydrogen infrastructure.
2. Adoption of policies promoting fuel cell technologies (Ball and Wietschel, 2009a; Kelly et al., 2010).

The development of an efficient infrastructure for producing and delivering hydrogen appears as a key factor, both to achieve the hydrogen transition and for its future development (NRC, 2004; DOE, 2009a; Hajimiragha et al., 2009). Thus, the design of an economically viable hydrogen Supply Chain (SC) could play an important and decisive role on the final market introduction and success of hydrogen as an alternative fuel and energy carrier. Despite several valuable contributions in the area, the dispute over the future of hydrogen is still in the air. The situation reveals in fact the need for more models and tools capable of assisting decision makers in this issue (DOE, 2008).

2.1.2 Hydrogen infrastructure optimization

In a pioneering work, Van den Heever and Grossmann (2003) focused on the integration of production planning and reactive scheduling in the optimization of a refinery hydrogen network. Here, hydrogen was part of a refinery production process and the problem was formulated as a deterministic multi-period mixed-integer non-linear program (MINLP). The SC design was out of the scope of the work, which was dedicated to integrate the mid-term planning decisions and reactive scheduling of a two-echelon single commodity network, for which valuable insights and solution strategies were devised. Then, Hugo et al. (2005) proposed a deterministic multi-period MILP approach for the long-term strategic planning of a multi-echelon hydrogen network taking into account economic and environmental concerns. Nearly at the same time Almansoori and Shah (2006) presented a steady-state deterministic mathematical programming framework to design and operate a future British hydrogen SC by optimizing the economical performance of the network. This work focused on getting an accurate and developed data set capable of producing reliable results for a generalized model built and presented in particular detail.

More recently significant contributions appeared from places where hydrogen has started to be seen as an important alternative fuel and energy carrier. Specifically, Ingason et al. (2008) proposed an MILP approach to find the most economical site selection for some hydrogen production technologies in Iceland. Lin et al. (2008b) devised an MILP model to optimize a hydrogen station sitting in Southern California by minimizing the fuel-travel-back time. Using the same framework, Lin et al. (2008a) proposed an approach to determine the least-cost hydrogen infrastructure design considering different technological alternatives to be established in the region of Southern California. These approaches provide a more detailed view of the operational level of hydrogen networks focused on the region under study.

Kim et al. (2008) developed a steady-state MILP model for the optimization of a hydrogen SC under demand uncertainty, which accounted for the optimization of the maximum, minimum and average scenarios. This work can be seen as the first stochastic approach to optimize a hydrogen network. Li et al. (2008) extended the previous work by Hugo et al. (2005) where the case study of China was specifically analyzed under a multi-period MILP framework for a hydrogen infrastructure design and optimization. In the same year, Kim and Moon (2008b) introduced a multi-objective optimization approach for the strategic design of a hydrogen infrastructure taking into account cost and safety. The mathematical formulation of this model is an extension of the work by Kim et al. (2008). Also Guillén-Gosálbez et al. (2010) have presented a new deterministic and multi-period MILP framework for hydrogen network optimization considering cost and environmental impact. In this work, they focus on the analysis of the environmental impact through a life cycle analysis perspective, and a new algorithm is presented to reduce the large computational burden associated to the problem resolution. Later, Kamarudin et al. (2009) developed a deterministic MILP single-period model to optimize a future hydrogen infrastructure in Malaysia, and Almansoori and Shah (2009) extended their previous snapshot formulation (Almansoori and Shah, 2006) to present a multi-period MILP model for optimizing the operation of a future hydrogen supply chain in Great Britain. Most recently, Guillén-Gosálbez and Grossmann (2010) extended their previous formulation to consider uncertainty in the environmental damage assessment model for their original bi-criterion environmentally conscious supply chain problem. Despite the theoretically sound capabilities of their novel approach, the framework relied on a joint chance-constraint reformulation that turned the stochastic model into a non-convex MINLP problem considerably more difficult to solve.

2.1.3 Hydrogen infrastructure optimization under uncertainty

Almost all the modeling approaches described above imply the use of deterministic models assuming that all problem parameters can be perfectly known in advance. Deterministic models are solved for the most likely scenario, thus neglecting any possible variability in its parameter values. This strategy provides solutions that perform well in the mean scenario, but can yield poor results for other possible values of the uncertain parameters. In practice, however, there are several sources of uncertainty that can affect the calculations (*e.g.*, prices of final products, operating cost, demand, resource availability, *etc.*). These uncertainties can be handled using stochastic models that incorporate them at the modeling stage. The use of stochastic programming allows to assess the SC decisions in the space of uncertain parameters, before the final solution is taken. In particular, the two prevalent stochastic approaches are stochastic programming with recourse (Linderoth et al., 2006) and robust optimization (Ben-Tal and Nemirovski, 1998).

The use of stochastic modeling tools has allowed for the incorporation of different sources of uncertainty into the decision-making process. Specifically, demand has been the most studied source of uncertainty (Sahinidis, 2004; Shapiro, 2004; Grossmann, 2004, 2005; Guillén-Gosálbez et al., 2005; Melo et al., 2009; Guillén-Gosálbez et al., 2006b,a), whereas other uncertainties, especially those appearing in the coefficients of the objective function (product prices, operating cost, *etc.*) have received much less attention. In the context of designing a hydrogen SC, the latter sources of uncertainty play a major role. This is due to the energy price variability to which the actual financial market is exposed, which recently reached historic levels and is largely dependent on fossil fuel prices.

Stochastic models that consider the variability of the uncertain parameters typically optimize the expected economic performance of the system. These approaches can lead to solutions that perform well on average but have large probabilities of unfavorable scenarios.

In practice, decision makers may have different attitudes towards the financial risk associated with the investment on a project under uncertainty. Many decision-makers tend to be risk averse, that is to say, they aim to avoid unfavorable situations thus showing a clear preference for solutions with lower variability for a given budget. The idea underlying risk management is the incorporation of the trade-off between risk and cost within the decision making process (Ferrer-Nadal et al., 2008; Guillén-Gosálbez et al., 2005; Barbaro and Bagajewicz, 2004). This leads to a multi-objective optimization problem in which the expected performance and a specific risk measure are the objectives considered. In this way, the concept of Supply Chain Management (SCM) coupled with risk management tools offer the opportunity of reducing the impact of unexpected events through an integrated multi-objective management of the network.

The aim of this work is to provide a mathematical programming framework for long-term design and planning of hydrogen supply chains for vehicle use under uncertainty in their economic performance with the ability to handle the financial risk associated to market changes.

Our approach is based on a novel multi-scenario MILP that accounts for the uncertainty associated with the coefficients of the objective function. The financial risk is explicitly measured *via* the worst case, which is appended to the objective function as an additional criterion to be optimized. The resulting large scale bi-criterion MILP tends to be computationally prohibitive as the number of time periods, alternatives included in the superstructure, potential locations and scenarios increases. Hence, our modeling framework is complemented by an efficient decomposition method that expedites the search of the Pareto solutions of the multi-objective model by exploiting its mathematical structure. The capabilities of the proposed modeling framework and solution strategy are illustrated through its application to a real case study based on Spain, for which valuable insights are obtained.

The chapter is organized as follows. Firstly, the problem studied is formally stated in Section 2.2 and the corresponding mathematical formulation is next provided in Section 2.3 . The decomposition strategy developed to solve the problem is introduced in Section 2.4 . In Section 2.5 a case study is described together with the capabilities of the proposed modeling framework and solution strategy. Finally, in Section 2.6 conclusions and remarks are presented.

2.2 Problem statement

The design problem addressed in this work has as objective to determine the optimal configuration of a three-echelon hydrogen supply chain for vehicle use (production-storage-market) in terms of minimizing the expected total discounted cost and the associated financial risk. The superstructure of the three-echelon SC taken as reference in this work is depicted in Figure 2.1. This network includes a set of plants, where hydrogen can be produced (pentagons), and a set of storage facilities (circles), where hydrogen is stored for being delivered to the final customers (rectangles) in order to meet their demand. The geographical framework includes the region of interest (*e.g.* country), which can be divided into a set of potential locations. These potential locations correspond to different sub-regions of the original region of interest, each one characterized by a given hydrogen demand. We consider that the set of potential locations of the problem along with the associated geographical distribution of the demand are input data to the problem. Consequently, the design problem can be formally stated as follows: Given are the hydrogen demands, fixed time horizon and number of time periods in which it is divided, the set of available production, storage and transportation technologies, the capacity limitations of plants and storage facilities, the costs associated with the network operation (production, transportation and inventory costs), the investment cost, the probabilistic information that describe the uncertain parameters (*i.e.*, type of probability distribution, number of scenarios, mean and variance) and interest rate. The goal is to determine:

1. The SC design: Involving the number, type, location and capacity of plants and storage facilities, along with the number and type of transportation units (*e.g.* tanker trucks, railway tube cars, *etc.*) and transportation links to be established between the potential locations;

2. the associated planning decisions: Including the production rates of plants, inventory levels at the storage facilities and flows of hydrogen between plants and storage facilities; in order to simultaneously minimize the expected total discounted cost and the associated financial risk of the network.

2.3 Mathematical model

The model presented in this work is inspired by previous approaches introduced by Almansoori and Shah (2006) , Almansoori and Shah (2009) Kim et al. (2008), Kim and Moon (2008b), Guillén-Gosálbez et al. (2010) and Guillén-Gosálbez and Grossmann (2010). Our approach presents on a multi-objective, multi-period and multi-scenario stochastic optimization model that extends the deterministic mathematical formulation previously presented by Guillén-Gosálbez et al. (2010) to account for different decision-making perspectives towards risk. More specifically, our mathematical formulation, which is based on the superstructure depicted in Figure 2.1, extends the previous formulations in order to account for a larger amount of production and transportation technologies. Particularly, the model considers the possibility of establishing different production and storage facilities in a set of potential locations with known demand and uncertain economic parameters. Financial risk management tools are introduced in the formulation to reduce the probability of occurrence of undesirable outcomes. Thereby,

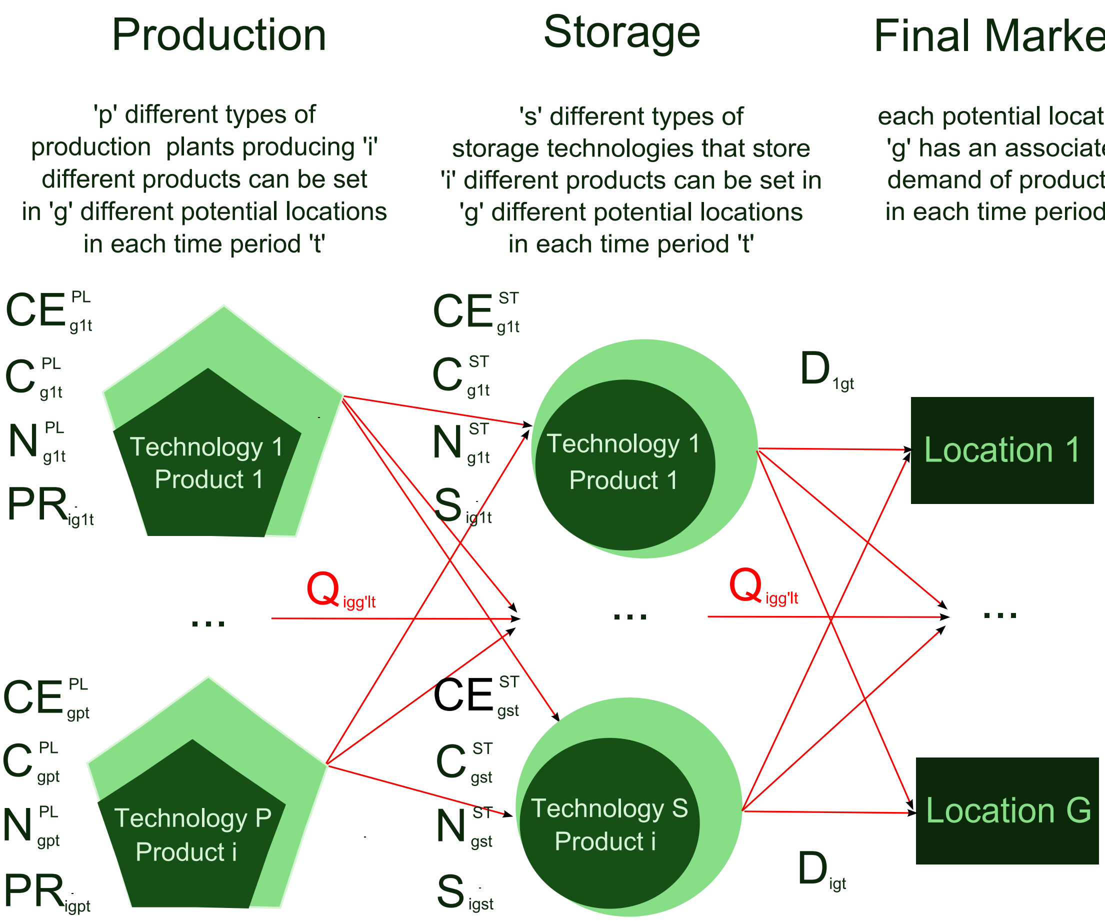


Figure 2.1: Three echelon supply chain taken as reference.

this model includes the financial risk, represented by the worst case, as an additional criterion to be optimized. For the sake of completeness of our work, we next provide a detailed description of the MILP model notation, constraints, and objective function equations.

2.3.1 Model features

The following assumptions are considered in our study:

- Demand scheme: The model is prepared to design a network capable of satisfying a given hydrogen demand pattern.
- Multi-period scheme: The network design covers a fixed time horizon divided in several periods of time accounting for prices and demand variations.
- Uncertainty: It is introduced in the operating costs of the network. The rest of the parameters involved are considered to be deterministic. The reason for that is to analyze separately the effects of different sources of uncertainty. However the model could be easily extended in order to account for other parameter uncertainties.
- Economies of scale: In order to introduce the effect of the economies of scale, while avoiding to introduce non-linearities in the problem formulation, we used the six-tenths-factor rule.
- Capacity expansions: In the model an initial demand is assumed, and the network design can be adapted to demand variation by introducing capacity expansions. However, in the case of pipelines it is assumed that they can only be constructed once in the life time of the project. With regard to the maritime transport, we consider that the ships are not purchased, but rather they are hired.

Also remark that aspects related with a detailed design and sizing of the facilities are at this point outside the scope of this work, which concentrates on the major SC strategic decisions rather than on the particular features of the individual SC echelons.

2.3.2 Mass balance constraints

The mass balance regulates the flow interactions along the supply chain, and must be satisfied for every hydrogen form i, liquid or gas, in each potential location g and time period t. Thus, for every location, the sum of the initial inventory S_{igst-1} plus the amount produced (PR_{igpt}) and the input flow rate ($Q_{ig'glt}$) must equal the final inventory (S_{igst}) plus the amount delivered to the customers (D_{igt}) and the output flow rate ($Q_{igg'lt}$) of hydrogen:

$$\begin{aligned} &\sum_{s\in SI(i)} S_{igst-1} + \sum_{p\in PI(i)} PR_{igpt} + \sum_{g'\neq g}\sum_{l\in LI(i)} Q_{ig'glt} \\ &= \sum_{s\in SI(i)} S_{igst} + D_{igt} + \sum_{g'\neq g}\sum_{l\in LI(i)} Q_{igg'lt} \qquad \forall i,g,t \end{aligned} \tag{1}$$

In Eq. (1), $SI(i)$ represents the set of storage technologies that can be used for product form i, $LI(i)$ is the set of transport modes that can transport product form i, and $PI(i)$ denotes the production facilities that can produce product form i.

2.3.3 Capacity constraints

2.3.3.1 Plants

In Eq. (2) the total production rate of hydrogen form i in location g produced *via* technology p in period t (PR_{igpt}) is restricted to be lower than the existing capacity of the plant and higher than a minimum desired percentage, τ, of the capacity installed. The capacity of each production technology p at location g in period t is represented by C^{PL}_{gpt}.

$$\tau C^{PL}_{gpt} \leq \sum_{i\in IP(p)} PR_{igpt} \leq C^{PL}_{gpt} \qquad \forall g,p,t \tag{2}$$

In this equation, $IP(p)$ denotes the set of hydrogen forms that can be produced by technology p. The capacity of each technology p in period t is calculated by adding the expansion in capacity (CE^{PL}_{gpt}) executed in period t and location g to the existing capacity at the end of the previous period:

$$C^{PL}_{gpt} = C^{PL}_{gpt-1} + CE^{PL}_{gpt} \qquad \forall g,p,t \tag{3}$$

Equation (4) is applied to limit capacity expansions within lower and upper bounds, obtained by multiplying the number of plants installed, which is denoted by the integer variable N^{PL}_{gpt} and the minimum and maximum capacities associated with each technology p ($\underline{PC^{PL}_p}$ and $\overline{PC^{PL}_p}$, respectively).

$$\underline{PC^{PL}_p} N^{PL}_{gpt} \leq CE^{PL}_{gpt} \leq \overline{PC^{PL}_p} N^{PL}_{gpt} \qquad \forall g,p,t \tag{4}$$

2.3.3.2 Storage facilities

In a similar way as for the production facilities, Eq. (5) enforces the total inventory of product in form i kept at the end of period t in the storage facilities of type s installed in location g (S_{igst}), to be lower than the available capacity. Here, C^{ST}_{gst} represents the storage capacity of product form i during period t in location g associated with storage technology s.

$$\sum_{i \in IS(s)} S_{igst} \leq C^{ST}_{gst} \qquad \forall g, s, t \tag{5}$$

This equation constrains the amount of hydrogen delivered from the storage facility to the customers to be lower than its capacity. In this equation, $IS(s)$ represents the set of product forms i that can be stored by technology s. This approach also uses the storage period θ, as previously introduced by Almansoori and Shah (2006), which is multiplied by two in order to cover fluctuations in both supply and demand as well as plant interruptions:

$$2(\theta D_{igt}) \leq \sum_{s \in SI(i)} C^{ST}_{gst} \qquad \forall i, g, t \tag{6}$$

In Eq. (6) $SI(i)$ denotes the set of storage technologies s that can handle product forms i. In a similar manner as occurred with the manufacturing plants, the capacity of the storage technology s at any time period t is determined from the previous one and the expansion in capacity executed in the same period:

$$C^{ST}_{gst} = C^{ST}_{gst-1} + CE^{ST}_{gst} \qquad \forall g, s, t \tag{7}$$

Finally, the value of the capacity expansion for storage facilities CE^{ST}_{gst} is bounded within lower and upper limits.

$$\underline{SC^{ST}_s} N^{ST}_{gst} \leq CE^{ST}_{gst} \leq \overline{SC^{ST}_s} N^{ST}_{gst} \qquad \forall g, s, t \tag{8}$$

2.3.4 Transportation constraints

In this block of equations, we introduce a binary variable $X_{gg'lt}$ to represent the existence or absence of a transportation link of type l (*e.g.*, tanker trucks, railway tube cars, *etc.*) between locations g and g' in time period t.

$$\begin{gathered} \underline{QC_{gg'l}} X_{gg'lt} \leq \sum_{i \in IL(l)} Q_{igg'lt} \leq \overline{QC_{gg'l}} X_{gg'lt} \\ \forall g, g'(g \neq g'), l \neq pipeline, t \end{gathered} \tag{9}$$

A zero value of the aforementioned binary variable prevents the flow of material that can be transported *via* transportation technology l between g and g' from taking place, whereas a value of one allows the transport within some lower $\underline{QC_{gg'l}}$ and upper bounds $\overline{QC_{gg'l}}$. In Eq. (9), $IL(l)$ denotes the set of product forms i that can be transported by transport mode l.

Eq. (10) is similar to Eq. (9), but applies only to pipelines. Specifically, we assume that if a pipeline is constructed, then the associated transportation link will remain open throughout the entire time horizon:

$$\sum_{t'<t+1} \underline{QC_{gg'l}} X_{gg'lt'} \leq \sum_{i \in IL(l)} (Q_{igg'lt} + Q_{ig'glt}) \leq \sum_{t'<t+1} \overline{QC_{gg'l}} X_{gg'lt'} \quad \forall g, g'(g \neq g'), l = pipeline, t \tag{10}$$

Furthermore, only one transportation link involving pipelines can be constructed at most during the entire time horizon:

$$\sum_{t'<t+1} X_{gg'lt'} \leq 1 \qquad \forall g, g'(g \neq g'), l = pipeline, t \tag{11}$$

We assume that a location can either import or export hydrogen, but not both at the same time. This is because if a location can only satisfy its needs by importing from other locations, it would not need to export to other locations:

$$X_{gg'lt} + X_{g'glt} \leq 1 \qquad \forall g, g'(g \neq g'), l, t \tag{12}$$

Specific constraints are appended to the model formulation to handle the case of maritime transportation devices. Hence, the binary variables $X_{gg'lt}$ denoting the existence of transportation links are forced to take a zero value in some particular cases *via* Eqs. 13 and (14), which prevent ships from transporting materials between locations with no harbors and also avoid the use of road transportation devices between harbors that are not connected by road transportation.

$$\begin{aligned} & X_{gg'lt} = 0 \qquad \forall l, g, g' \in LG \\ & LG = \{l, g, g' : (l = ship) \wedge ((g, g') \notin SGG(gg'))\} \end{aligned} \tag{13}$$

$$\begin{aligned} & X_{gg'lt} = 0 \qquad \forall l, g, g' \in LG' \\ & LG' = \{l, g, g' : (l \neq ship, pipeline) \wedge ((g, g') \in SGG'(gg'))\} \end{aligned} \tag{14}$$

In these constraints, $SGG(g, g')$ is the subset of allowable maritime links, whereas $SGG'(g, g')$ is the subset of maritime links (*i.e.*, $SGG'(g, g') \subset SGG(g, g')$) that cannot be connected through road transportation units.

2.3.5 Demand satisfaction constraint

The total amount of hydrogen consumed (D_{igt}) is constrained to be lower than the total hydrogen demand ($\overline{D_{gt}}$) in each location and period, and higher than a given minimum demand satisfaction level ($dsat$) :

$$\overline{D_{gt}}dsat \leq \sum_{i} D_{igt} \leq \overline{D_{gt}} \qquad \forall g, t \tag{15}$$

2.3.6 Objective function equations

The model presented considers that the coefficients of the objective function (*e.g.* facility investment and variable costs and transportation costs) are uncertain and that their variability can be described through a set of scenarios with given probability of occurrence. As a result, the cost associated with the establishment and operation of the SC is not a single nominal value, instead it is a stochastic variable that follows a discrete probability function. In this context, the optimization method must identify the set of solutions (*i.e.*, strategic SC decisions) that simultaneously minimize the expected value of the cost distribution as well as its financial risk level. The main advantage of the scenario-based approach is that it allows dealing with any type of probability function. Furthermore, this approach avoids the non-linearities associated with the reformulation of the probabilistic constraints used in robust optimization (Ben-Tal and Nemirovski, 1998; Janak et al., 2007; Guillén-Gosálbez and Grossmann, 2009).

2.3.6.1 Expected cost

The expected total cost is given by the mean value of the discrete distribution of the cost associated to each scenario realization:

$$E[TDC] = \sum_{e} prob_e TDC_e \tag{16}$$

The total discounted cost attained in each particular scenario realization (TDC_e) is calculated as the sum of the discounted costs associated with each time period t:

$$TDC_e = \sum_{t} \frac{TC_{te}}{(1+ir)^{t-1}} \qquad \forall e \tag{17}$$

In the previous expressions, e is a subscript that represents a particular scenario e and $prob_e$ is the probability of occurrence of this scenario. In Eq. (17), ir represents the interest rate and TC_{te} is the total amount of money spent in period t and scenario e, which includes the capital (FCC_t, TCC_t) as well as operating costs (FOC_{te}, TOC_{te}) given by the production, storage and transportation facilities of the network:

$$TC_{te} = FCC_t + FOC_{te} + TCC_t + TOC_{te} \qquad \forall t, e \tag{18}$$

In practice, it will be possible to have a good estimate of the capital cost at the design stage, since it is usually agreed before the establishment of a new facility. On the other hand, the value of the operating cost will fluctuate according to the market trends. Hence, in Eq. (17) FCC_t and TCC_t can be regarded as non-scenario dependent variables, whereas FOC_te and TOC_te will in general depend on the specific scenario realization.

2.3.6.2 Facility capital cost

The facility capital cost over period t (FCC_t) is determined from the capacity expansions made in the manufacturing plants and storage facilities during that period:

$$FCC_t = \sum_g \sum_p \left(\alpha^{PL}_{gpt} N^{PL}_{gpt} + \beta^{PL}_{gpt} CE^{PL}_{gpt} \right) \\ + \sum_g \sum_s \left(\alpha^{ST}_{gst} N^{ST}_{gst} + \beta^{ST}_{gst} CE^{ST}_{gst} \right) \qquad \forall t \tag{19}$$

The parameters, α^{PL}_{gpt}, β^{PL}_{gpt}, α^{ST}_{gst} and β^{ST}_{gst} are the fixed and variable investment terms corresponding to plants and storage facilities, respectively. These parameters reflect the concept of economies of scale.

2.3.6.3 Facility operating cost

The facility operating cost term is obtained by multiplying the unit production and storage costs (upc_{igpte} and usc_{igste}, respectively), which are regarded as uncertain parameters, with the corresponding production rates and average inventory levels:

$$FOC_{te} = \sum_i \sum_g \sum_{p \in PI(i)} upc_{igpte} PR_{igpt} \\ + \sum_i \sum_g \sum_{s \in SI(i)} usc_{igste} \left(\theta D_{igt} \right) \qquad \forall t, e \tag{20}$$

2.3.6.4 Transportation capital cost

The transportation capital cost, which includes the cost of the trucks and railcars is calculated *via* Eq. (21):

$$TCC_t = \sum_{l \neq ship, pipeline} N^{TR}_{lt} \cdot cc_{lt} + PCC_t \qquad \forall t \tag{21}$$

Here, PCC_t is the pipeline capital costs, cc_{lt} represents the capital cost associated with transport mode l in period t, and N^{TR}_{lt} is an integer variable that denotes the total number of transportation units of type l purchased in period t. Note that ships and pipelines are excluded from the first term of the summation. The reason for this is that the model assumes that ships are hired for carrying out the specific transportation tasks (*e.g.*, outsourcing), whereas the capital cost of pipelines is calculated *via* the following equation:

$$PCC_t = \sum_{g} \sum_{g' \neq g} \sum_{l=pipeline} upcc_t X_{gg'lt} distance_{gg'} \qquad \forall t \tag{22}$$

where $upcc_t$ is the unit capital cost of the pipeline per unit of length built and $distance_{gg'}$ denotes the distance between potential locations g and g'.

The average number of trucks and/or railcars required to satisfy a certain flow between different locations is computed from the flow rate of products between the locations ($Q_{igg'lt}$), the transportation mode availability (av_l), the capacity of a transport container ($tcap_l$), the average distance traveled between the locations ($distance_{gg'}$), the average speed ($speed_l$) and the loading/unloading time ($lutime_l$), as stated in Eq. (23):

$$\sum_{t'<t+1} N_{lt'}^{TR} \geq \sum_{i \in IL(l)} \sum_{g} \sum_{g' \neq g} \sum_{t} \frac{Q_{igg'lt}}{av_l tcap_l} \left(\frac{2distance_{gg'}}{speed_l} + lutime_l \right) \qquad \forall l \neq ship, pipeline \tag{23}$$

The total number of transportation units available in any period t includes those purchased in the same period t as well as those acquired in previous periods t'. Therefore, the left hand side of the inequality in Eq. (23) represents the summation of all the transportation units purchased in all the time periods t' up to the actual period t (*i.e.*, $t' = t$). Also here, $IL(l)$ denotes the set of product forms i that can be transported by transport mode l. For the purpose of simplicity, this work assumes that each transportation facility can only operate between two predefined locations. For this reason the distance between locations g and g' ($distance_{gg'}$) is multiplied by two, so the model accounts for the return journey of the trucks/railcars.

2.3.6.5 Transportation operating cost

The total operating cost associated with the transportation tasks carried out in scenario e in period t (TOC_{te}) is determined from Eq. (24):

$$TOC_{te} = ROC_{te} + POC_{te} + SOC_{te} \qquad \forall t, e \tag{24}$$

where ROC_{te}, POC_{te} and SOC_{te} are the operating costs associated with road transportation technologies and railway, pipelines and ships, respectively. The first term includes the fuel (FC_{te}), labor (LC_{te}), maintenance (MC_{te}) and general costs (GC_{te}):

$$ROC_{te} = FC_{te} + LC_{te} + MC_{te} + GC_{te} \qquad \forall t, e \tag{25}$$

The fuel cost is a function of the fuel price ($fuelp_{lte}$) and fuel consumption:

$$FC_{te} = \sum_{g} \sum_{g' \neq g} \sum_{l \neq ship,pipeline} \sum_{i \in IL(l)} fuelp_{lte} \frac{2distance_{gg'} Q_{igg'lt}}{fuelc_l tcap_l} \qquad \forall t, e \tag{26}$$

Note that the main source of uncertainty here is the fuel price, since it cannot be perfectly known in advance at the design stage. In Eq. (26), the fractional term represents the fuel usage, and it is determined from the total distance traveled in a trip (2 $distance_{gg'}$), the fuel consumption of transport mode l ($fuelc_l$) and the number of trips made per period of time ($\frac{Q_{igg'lt}}{tcap_l}$). The labor transportation cost is described as a function of the driver wage in scenario e ($wage_{lte}$) and total delivery time (*i.e.*, the term inside the brackets):

$$\begin{aligned} LC_{te} = & \sum_{g} \sum_{g' \neq g} \sum_{l \neq ship,pipeline} \sum_{i \in IL(l)} wage_{lte} \\ & \times \left[\frac{Q_{igg'lt}}{tcap_l} \left(\frac{2distance_{gg'}}{speed_l} + lutime_l \right) \right] \qquad \forall t, e \end{aligned} \tag{27}$$

The maintenance cost, which accounts for the general maintenance of the transportation systems, is a function of the cost per unit of distance traveled in scenario e (cud_{lte}) and total distance driven:

$$MC_{te} = \sum_{g} \sum_{g' \neq g} \sum_{l \neq ship,pipeline} \sum_{i \in IL(l)} cud_{lte} \frac{2distance_{gg'} Q_{igg'lt}}{tcap_l} \qquad \forall t, e \tag{28}$$

The general cost includes the transportation insurance, license and registration, and outstanding finances. It can be determined from the unit general expenses in scenario e (ge_{lte}) and number of transportation units as follows:

$$GC_{te} = \sum_{l \neq ship,pipeline} \sum_{t' \leq t} ge_{lte} N^{TR}_{lt'} \qquad \forall t, e \tag{29}$$

Equation (30) determines the pipeline operating costs from the unit operating cost of the pipelines in scenario e ($upoc_{te}$) and the freight to be delivered.

$$POC_{te} = \sum_{g} \sum_{g' \neq g} \sum_{l = pipeline} \sum_{i \in IL(l)} upoc_{te} Q_{igg'lt} \qquad \forall t, e \tag{30}$$

Finally, Eq. (31) calculates the ship operating costs based on the unit operating costs for maritime transportation in scenario e ($usoc_{te}$), the time required to deliver the hydrogen and the cargo:

$$SOC_{te} = \sum_{g} \sum_{g' \neq g} \sum_{l = ship} \sum_{i \in IL(l)} usoc_{te} \left(distance_{gg'} Q_{igg'lt} \right) \qquad \forall t, e \tag{31}$$

2.3.6.6 Risk management

The traditional approach to address optimization under uncertainty relies on formulating a single-objective optimization problem where the expected performance of the system is the objective to be optimized. This strategy does not allow controlling the variability of the objective function in the uncertain space. In other words, optimizing the expected economic performance of a SC does not imply that the process will yield better results at a certain level considering the whole cost distribution. The underlying idea in risk management is to incorporate the trade-off between financial risk and expected cost into the decision-making procedure. This gives rise to a multi-objective optimization problem in which the expected performance and a specific risk measure are the objectives considered. The solution of such a problem is given by a set of Pareto solutions that represent the optimal trade-off between expected performance and risk level. In mathematical terms, the financial risk associated with a design project can be defined as the probability of not meeting a certain target profit (maximization) or exceeding a certain cost level (minimization) (Barbaro and Bagajewicz, 2004).

Hence, the financial risk associated with a design x and a cost target Ω can be expressed as follows:

$$\mathrm{Risk(x,\Omega) = P[TDC(x) \geq \Omega]} \tag{32}$$

In Eq. (32) $TDC(x)$ is the total cost associated to a hydrogen network design x, that is, the cost resulting after the uncertainty has been unveiled and a scenario has been realized.

Specifically in this work, the probability of meeting unfavorable scenarios is controlled by adding the worst case cost as an additional objective to be minimized. This metric is easy to implement and shows good numerical performance in stochastic models, as pointed out by Bonfill et al. (2004), while at the same time circumvents the introduction of more binary variables in the formulation. The worst case value can be easily determined from the maximum cost attained over all scenarios:

$$\mathrm{WC \geq TDC_e} \qquad \forall \mathrm{e} \tag{33}$$

The inclusion of the worst case as an alternative objective to be minimized along with the expected total cost leads to the following bi-criterion MILP formulation:

$$\begin{aligned} (MOP) \qquad & \min_{x,y,z} \left(E[TDC](x,y,z), WC(x,y,z)\right) \\ s.t. \qquad & Eqs.\ (1) - \ (33) \\ & x \in \Re,\ y \in \{0,1\},\ z \in N \end{aligned} \tag{34}$$

where x, y and z denote the continuous, binary and integer variables of the problem, respectively. The aforementioned multi-objective problem can be solved by standard algorithms for multi-objective optimization such as the ϵ-constraint or the weighted-sum method (Ehrgott, 2000b). The weighted-sum method is only rigorous for the case of convex Pareto sets, whereas the ϵ-constraint method is also rigorous for the non-convex case, which turns out to be our case. This method entails solving a set of instances of problem (P) corresponding to different values of the auxiliary parameter ϵ:

$$
\begin{aligned}
(P) \qquad & \min_{x,y,z} E[TDC](x,y,z) \\
s.t. \qquad & Eqs.\ 1 - \ 33 \\
& WC(x,y,z) \leq \epsilon \\
& \underline{\epsilon} \leq \epsilon \leq \overline{\epsilon} \\
& x \in \Re,\ y \in \{0,1\},\ z \in N
\end{aligned}
\tag{35}
$$

where the lower and upper bounds within which the epsilon parameter must fall ($i.e.$,$\epsilon \in [\underline{\epsilon}, \overline{\epsilon}]$) are obtained from the optimization of each scalar objective separately.

2.4 Solution approach: two-step sequential approach

Model (P) tends to be computationally intensive, especially as the number of potential locations, available technologies and time periods increases. To circumvent this issue, we introduce next a decomposition strategy that expedites its solution and makes it possible to address medium/large size instances of (P) that might appear when dealing with realistic problems found in practice.

Our solution method is a two-step sequential approach that exploits the specific structure of the model. This solution procedure is inspired on previous bi-level decomposition methods presented so far in the literature (Iyer and Grossmann, 1998; Guillén-Gosálbez et al., 2010). The method exploits the fact that in practical problems the relaxation of the integer variables of the full space model tends to be very tight. In other words, the solution that is obtained when (P) is solved defining z as continuous variables instead as integers is indeed very close to the true optimal solution of (P). The reason for this is that in practice these integer variables take large values, since they represent the number of facilities to be established in big regions that cover high demands. On the other hand, the binary relaxation tends to be much weaker.

Based on this observation, we propose a method to solve (P) that relies on decomposing it into two hierarchical levels: a lower relaxed master problem (MP) and an upper bounding slave problem (SP), that are solved in a sequential way.

2.4.1 Master problem

The master problem (MP) is obtained from model (P) as follows. The integer variables z (*i.e.*, N^{PL}_{gpt}, N^{ST}_{gst} and N^{TR}_{lt}) representing the number of production, storage and transportation technologies, are relaxed by reformulating them as continuous variables. Note that by doing so, the solution of the model is highly expedited, since the combinatorial complexity associated with the presence of integer variables is reduced to a large extent. The master problem can therefore be expressed as follows.

$$\begin{aligned} (MP) \qquad & \min_{x,y,rz} E[TDC](x,y,rz) \\ s.t. \qquad & Eqs.\ 1 - \ 33 \\ & WC(x,y,rz) \leq \epsilon \\ & \underline{\epsilon} \leq \epsilon \leq \overline{\epsilon} \\ & x, z \in \Re,\ y \in \{0,1\} \end{aligned} \tag{36}$$

where rz denotes the set of integer variables that are relaxed into continuous ones. The master problem is therefore a relaxation of the original problem, and for this reason it provides a lower bound on its global solution.

2.4.2 Slave problem

The upper bounding slave model (SP) corresponds to the same original full-space formulation in which the integer variables are obtained by rounding up (or down) the optimal values of the relaxed integer variables provided by the master problem solution. The solution of the slave problem provides an upper bound to the true optimal solution of (P), since its search space is contained in that of the full-space model. The slave problem has the following form:

$$\begin{aligned} (SP) \qquad & \min_{x,y,z} E[TDC](x,y,z) \\ s.t. \qquad & Eqs.\ 1\ 33 \\ & z = \lceil rz^* \rceil \\ & WC(x,y,z) \leq \epsilon \\ & \underline{\epsilon} \leq \epsilon \leq \overline{\epsilon} \\ & x \in \Re,\ y \in \{0,1\},\ z \in N \end{aligned} \tag{37}$$

where rz^* denotes the optimal values of the relaxed integer variables calculated by the master problem.

2.4.3 Remarks

- The model presented in this work can be easily extended in order to address uncertainty in the demand. This could be accomplished by reformulating some variables of the problem as scenario-dependent ones.
- As shown in the work by Gebreslassie et al. (2009b), minimizing the expected cost of a stochastic problem with uncertain coefficients in the objective function is equivalent to minimizing the cost in the mean scenario. Thus, in this particular case, single-objective stochastic models that attempt to optimize the expected performance of the network provide the same solution as deterministic approaches. This does not happen when risk management metrics are considered, since such consideration leads to a multi-objective problem.
- In practice, the decomposition strategy works better when the number of transportation units is left free in the slave problem. This occurs because the master problem tends to provide solutions that underestimate the requirements of transportation units, since the relaxation of the integer variables denoting the number of plants take more advantage of the concept of economies of scale, thus decreasing the transportation needs.
- The slave problem of our method might be infeasible for some values of epsilon. This might occur when the master problem will predict too optimistic solutions (*i.e.*, low worst case values) that cannot be attained in practice when the integrality requirement is enforced. This limitation can be overcome by removing the ϵ-constraint from the slave problem. As it will be discussed later in Section 2.5, this will result in a shift of the Pareto curve corresponding to the slave solution with respect to that provided by the master problem.
- The solution method presented in this chapter can be applied to previous optimization models for optimizing hydrogen SCs, either deterministic or stochastic, presented so far in the literature (Hugo et al., 2005; Almansoori and Shah, 2006; Li et al., 2008; Kim and Moon, 2008b; Kim et al., 2008). Our method might improve the practical implementation of these models, thus allowing for their application to more complex problems.

2.5 Case study based on Spain

The capabilities of our modeling framework and solution strategy are illustrated through their application to a case study based on a real-world problem. Specifically, the optimal design of a hydrogen SC for vehicle use in Spain is addressed. A superstructure is postulated following the scheme presented in Section 2.2, where all different alternatives for hydrogen production, storage and transportation are embedded.

The region of interest is Spain and the potential locations, which are depicted in Figure 2.2a, are its autonomous communities (*i.e.*, states of the country). Each of these regions has an associated hydrogen demand that can be fulfilled either locally or by importing hydrogen from other locations.

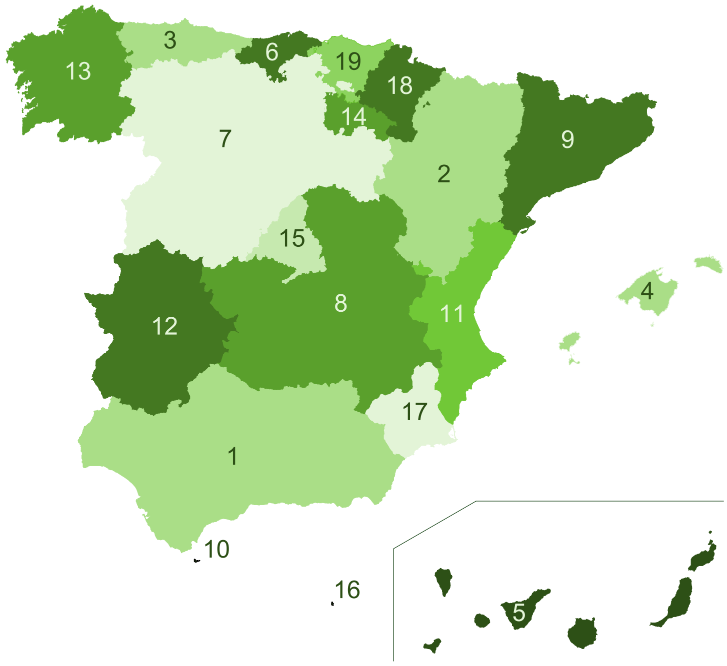
Figure 2.2b shows the alternatives considered for the production, storage and delivery of hydrogen. In this case, hydrogen (liquid or gas) can be produced *via* steam methane reforming, coal gasification, biomass gasification and wind electrolysis. Moreover, different storage and transportation facilities are considered for the case of liquid and gaseous hydrogen. With regard to the transportation technologies, we consider trucks, trains, ships and pipelines. In order to reflect the current increasing concerns and policies for greenhouse gas mitigation we have assumed that steam methane reforming and coal gasification facilities must include carbon sequestration devices.

The future expected hydrogen demand in Spain (see Table2.1) was calculated based on the annual report for gasoline and diesel fuel demands published by MITYC (2008). This information was translated into the corresponding hydrogen demand by using low calorific values provided by DOE (2009b). According to the report published by OSE (2009), the transportation sector accounts for 60% of the fuels consumption in Spain. A minimum demand satisfaction level of 90% was considered for the network calculations.

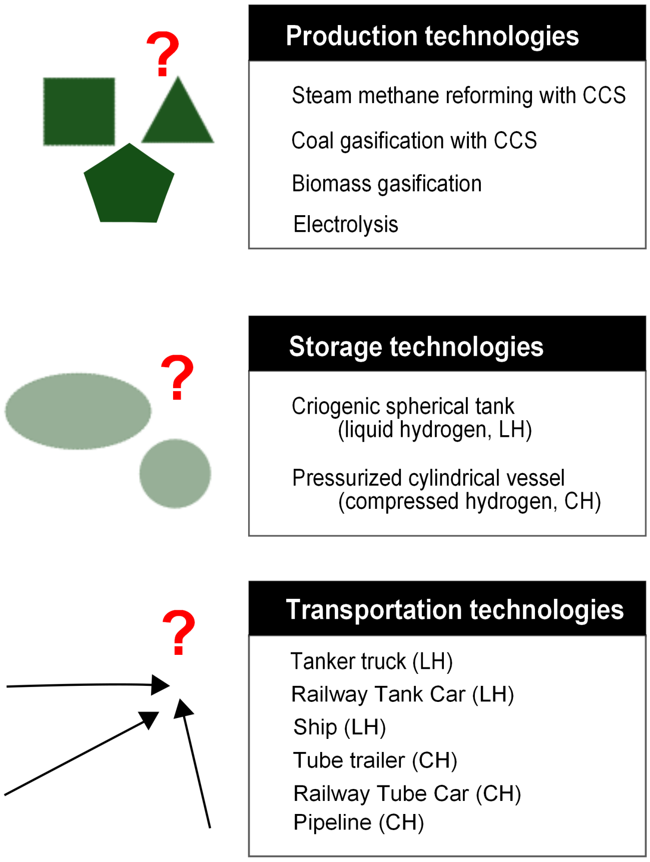
2.5.1 Transportation data set

All the data taken from sources before 2009 were actualized to current values by using the IPC factor (INE, 2009; IMF, 2009). The original data for United Kingdom (Almansoori and Shah, 2006) were adapted to Spain considering a correcting factor of 0.8, mainly based on an average factor relating the currencies of the two countries(IMF, 2009). To determine the pipelines capital cost, we follow the work by Amos (1998).

The maximum flow rate of liquid and compressed gaseous hydrogen transported *via* trucks or railcars $\overline{QC_{gg'l}}$ was assumed to be 960 000 $kg\ d^{-1}$, while the minimum flow rate $\underline{QC_{gg'l}}$ was assumed to be equal to the capacity of each transportation mode as done by Almansoori and Shah (2006). The capital and operating costs of the transportation modes are shown in Table 2.2. Particularly ship transportation costs were provided by Transmar S. L. (2009). The average delivery distances between the different potential locations were estimated taking as central points the capitals of the autonomous communities (see Table 2.3).



(a) Set of potential locations for the supply chain entities of the case study.



(b) Set of available production, storage and transportation technologies of the case study.

Figure 2.2: Definition of the case study.

Table 2.1: Hydrogen demand for $t = 1$ (assuming an annual increment of 5 %).

Potential locations	Indentifier	Demand (ton/d)
Andalucía	1	3,487
Aragón	2	937
Asturias	3	513
Baleares	4	695
Canarias	5	845
Cantabria	6	295
Castilla y León	7	2,131
Castilla la Mancha	8	1,509
Cataluña	9	3,372
Ceuta	10	20
C. Valenciana	11	2,210
Extremadura	12	599
Galicia	13	1,595
La Rioja	14	182
Madrid	15	2,200
Melilla	16	12
Murcia	17	836
Navarra	18	526
País Vasco	19	1,209

Table 2.2: Transportation costs

Transport data	Source	Tanker truck	Tube trailer	Tank railcar	Tube railcar	Ship[a]	Pipeline
Fuel economy (km/L fuel)	EcoTransIT (2009)	3.58	3.58	10.13	10.13	18.75	-
Fuel price (€/L fuel)	citeH2Demand	1.01	1.01	0.24	0.24	1.01	-
Driver wages (€/h)	Yang and Ogden (2007)	19.97	19.97	19.97	19.97	19.97	-
Maintenance expenses(€/km)	Almansoori and Shah (2006)	0.085	0.085	0.05	0.05	0.085	-
General expenses(€/day)	Almansoori and Shah (2006)	7.14	7.14	5.95	5.95	7.14	-
Loading/unloading time (h/trip)	Almansoori and Shah (2006)	2	2	12	12	48	-
Average speed (km/h)	EcoTransIT (2009)	55	55	45	45	16	-
Capacity (kg/trip)	Almansoori and Shah (2006)	4,082	181	9,072	454	4,082	-
Capital cost (€)	Almansoori and Shah (2006)	434,236	217,118	434,236	260,541		708,673[b]
	Kim et al. (2008)						
	Yang and Ogden (2007)						
	Amos (1998)						
Pipeline operating cost (€/kg day)	Kim et al. (2008)	-	-	-	-	-	0.05767
	Yang and Ogden (2007)						
	Amos (1998)						
Ship operating cost(€/h kg)	Transmar S. L. (2009)	-	-	-	-	0.00115	-
Minimum capacity (kg/day)	Almansoori and Shah (2006)	4,082	181	9,072	454	4,082	10
Maximum capacity (kg/day)	Almansoori and Shah (2006)	960,000	960,000	960,000	960,000	960,000	960,000

[a] Ship costs were obtained thanks to the collaboration with Transmar Ship & Forwarding Agency S.L.
[b] Pipeline capital cost depends on the delivery distance to be constructed and is expressed in €/km.

2.5.2 Production data set

Table 2.4 includes the production data set cost. The minimum and maximum production capacities for each plant type ($\underline{PC_p^{PL}}$ and $\overline{PC_p^{PL}}$) were assumed to be 10,000 and 480,000 kilograms per day ($kg\ d^{-1}$) respectively, based on commercial and near commercial medium to large hydrogen plants (Almansoori and Shah, 2006).

In this example, we devote our attention to the uncertainty associated with the unit production costs, and assume that the other operating parameters can be perfectly known in advance. We also consider that the uncertain values follow normal distributions. The mean values of these Gaussian bells from which the representative scenarios of the uncertain space are generated are given also in Table 2.4. A variance of 25% was assumed for steam methane reforming, (SMR), 5% for coal gasification (CG), and 10% for biomass gasification (BG) and electrolysis (E). The uncertain parameters were described by 50 scenarios using Monte Carlo sampling. The number of scenarios required to provide a solution within a specified confidence interval was determined by using the statistic-based methodology proposed by Law and Kelton (2000). Particularly in our case study, with a sample composed by 50 scenarios we can ensure a relative error of 0.5% and a confidence interval of 99.9% for the value of the $E[TDC]$ obtained.

The variability of the natural gas price affects the operating cost of steam methane reforming. To account for this variation, an increment of 20% on the original mean value from 2004 of hydrogen unit production costs was assumed for $t \leq 4$, while for $t \geq 5$ an increment of 25% was considered based on historic data provided by INE (2009) and NRC (2009).

The investment cost of carbon sequestration technologies (CCS) represents the 5% of the total capital investment cost in the case of SMR and 10% for CG (Ball and Wietschel, 2009b). For biomass gasification, no carbon sequestration technology was assumed, since it is regarded as a life cycle zero-carbon emission technology.

2.5.2.1 Storage data set

Table 2.5, available as supplementary material, includes the storage data set cost. The minimum and maximum storage capacities of each storage type ($\underline{SC_s^{ST}}$ and $\overline{SC_s^{ST}}$) were assumed to be 10,000 and 540,000 kilograms, respectively (Almansoori and Shah, 2006). The storage period (θ) is 10 days.

Table 2.3: Distances between sub-regions, km

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0	854	852	943	1,550	842	586	493	998	223	658	188	958	919	528	440	534	981	884
2	854	0	608	574	2,403	401	438	414	307	1,021	312	668	838	176	324	968	579	180	258
3	852	608	0	1,177	2,371	194	294	522	910	1,044	804	640	335	440	447	1,177	851	502	406
4	943	574	1,177	0	2,448	970	1,007	724	267	984	355	1,040	1,407	745	699	811	422	748	827
5	1,550	2,403	2,371	2,448	0	2,389	2,133	2,024	2,541	1,586	2,201	1,735	2,505	2,466	2,075	1,981	2,036	2,528	2,431
6	842	401	194	970	2,389	0	249	534	704	1,062	709	658	500	234	452	1,189	862	254	162
7	586	438	294	1,007	2,133	249	0	287	739	806	569	402	448	269	212	943	616	331	235
8	493	414	522	724	2,024	534	287	0	715	650	383	304	677	477	90	680	393	539	442
9	998	307	910	267	2,541	704	739	715	0	1,181	352	968	1,140	477	623	1,008	619	481	559
10	223	1,021	1,044	984	1,586	1,062	806	650	1,181	0	821	408	1,178	1,090	706	395	572	1,142	1,056
11	658	312	804	355	2,201	709	569	383	352	821	0	698	967	491	357	648	259	494	573
12	188	668	640	1,040	1,735	658	402	304	968	408	698	0	785	731	340	623	725	793	696
13	958	838	335	1,407	2,505	500	448	677	1,140	1,178	967	785	0	669	600	1,398	1,004	731	634
14	919	176	440	745	2,466	234	269	477	477	1,090	491	731	669	0	385	1,141	752	86	90
15	528	324	447	699	2,075	452	212	90	623	706	357	340	600	385	0	737	404	448	351
16	440	968	1,177	811	1,981	1,189	943	680	1,008	395	648	623	1,398	1,141	737	0	399	1,145	1,085
17	534	579	851	422	2,036	862	616	393	619	572	259	725	1,004	752	404	399	0	755	760
18	981	180	502	748	2,528	254	331	539	481	1,142	494	793	731	86	448	1,145	755	0	99
19	884	258	406	827	2,431	162	235	442	559	1,056	573	696	634	90	351	1,085	760	99	0

Table 2.4: Production costs

Production costs	Product	SMR	CG	BG	E
Capital cost (million €)[a]	Gaseous hydrogen	329	670	788	47
Unit production cost (€/kg)[b]	Gaseous hydrogen	0.82	0.92	1.49	2.28
Capital cost (million €)[a]	Liquid hydrogen	465	832	1,226	97
Unit production cost (€/kg)[b]	Liquid hydrogen	1.33	1.49	2.67	2.93

[a] Capital cost associated to a design production capacity of 480,000 kg /dsay for SMR, CG and BG, and of 50,000 kg/day for E. The design storage capacity is 540,000 kg.

[b] Uncertain parameter average value.

Table 2.5: Storage costs

Storage data	Product	Container type[a]	Cost
Capital cost (million €)	Liquid hydrogen	CST	106
Unit storage cost (€/kg day)	Liquid hydrogen	CST	0.0043
Capital cost (million €)	Gaseous hydrogen	PCV	1,645
Unit storage cost (€/kg day)	Gaseous hydrogen	PCV	0.0660

[a] Pressurized cylindrical vessel (PCV) for gaseous hydrogen and cryogenic spherical tank (CST) for liquid hydrogen

2.5.3 Results and discussion

2.5.3.1 Computational performance

We first solved several problems with various complexity levels. Particularly, we solved several instances of the problem considering different number of time periods with a length of two years each. The goal is to illustrate the performance of the proposed solution method as compared to the full-space model (*i.e.*, problem (P) without decomposition, relaxation or approximations).

All the problems were implemented in GAMS (Brooke et al., 1998) and solved in an AMD Phenom ^{TM} 8,600 B Triple-Core 2,300 MHz processor machine using CPLEX 9.0 as MILP solver. Tables 2.6, 2.7, 2.8 and 2.9 show the problem sizes and solution times for the proposed two-step sequential approach, solving each sub-problem to global optimality (*i.e.*, 0% gap), and the full-space method considering an optimality gap of 1%. This allows for a fair comparison between both methods as the sequential approach provides always solutions with less than 1% of optimality gap. We should note that in the slave problems of the algorithm, all the integer variables are relaxed except those denoting the number of transport units.

Due to space limitations, in the tables we only show the results corresponding to the extreme solutions of the Pareto set (*i.e.*, the minimum expected total discounted cost and the minimum worst case). Notice that in the case of the sequential approach, the quality of the solution shown in the table represents the difference between the solutions of the higher level and lower level problems, and not the optimality gap with which the sub-problems were solved.

As can be observed, for small problems (*i.e.*, $t = 2$), the full-space method is almost as efficient as the decomposition strategy, since the number of integer variables is very small. On the other hand, as the size of the problem increases the differences in CPU time are more significant. Specifically, for $t = 3$, the two-step sequential approach provides near optimal solutions (*i.e.*, solutions with an optimality gap lower than 1%) in CPU times that are on average one order of magnitude lower than those reported by the full-space approach.

Let us also note that the upper bounding (or slave) problem can be solved very quickly, whereas the lower level (or master) problem is the bottleneck of the proposed method. As shown in Table 2.6 to 2.9, the latter formulation decreases significantly the number of integer variables in comparison with the full space model. This reduces the combinatorial complexity of the problem to a large extent, thus alleviating its computational burden. It can also be observed that minimizing the worst case is computationally more expensive than minimizing the cost. Remark that the full space method is only able to minimize the worst case up to six time periods, whereas the two-step sequential approach can provide solutions for up to eight periods.

Table 2.6: Computational results for minimum worst case ($t \leq 4$).

	Variables						
	Binary	Integer	Continuous	Equations	Time (s)	WC(€)[a]	Quality (%)[b]
t = 1							
Full space	2,012	194	3,204	7,739	2.73	9.43×10^{11}	1
Two step					0.62		0.20
Master	2,012	–	3,398	7,739	0.48	9.42×10^{11}	
Slave	2,012	4	3,393	7,689	0.14	9.43×10^{11}	
t = 2							
Full space	4,024	388	6,356	15,377	9.48	1.07×10^{12}	1
Two step					2.36		0.30
Master	4,024	–	6,744	15,377	1.86	1.06×10^{12}	
Slave	4,024	8	6,735	15,327	0.50	1.07×10^{12}	
t = 3							
Full space	6,036	582	9,508	23,015	377.20	1.19×10^{12}	1
Two step					24.27		0.35
Master	6,036	–	10,090	23,015	19.97	1.18×10^{12}	
Slave	6,036	12	10,077	22,965	4.30	1.19×10^{12}	
t = 4							
Full space	8,048	776	12,660	30,653	228.47	1.32×10^{12}	1
Two step					58.70		0.51
Master	8,048	–	13,436	30,653	53.69	1.3×10^{12}	
Slave	8,048	16	13,419	30,603	5.01	1.32×10^{12}	

[a] Objective function values obtained by solving the full-space problem with 1 % optimality gap and the sub-problems of the sequential approach with 0% optimality gap

[b] Solution quality means here the optimality gap (%) of the solution provided by CPLEX (full-space problem) or the gap between the master and slave problems in the sequential approach

Table 2.7: Computational results for minimum worst case ($t \geq 5$).

	Variables						
	Binary	Integer	Continuous	Equations	Time (s)	WC(€)[a]	Quality (%)[b]
$t = 5$							
Full space	10,060	970	15,812	38,291	649.50	1.44×10^{12}	1
Two step					117.19		0.59
Master	10,060	–	16,782	38,291	113.97	1.43×10^{12}	
Slave	10,060	20	16,761	38,241	3.22	1.44×10^{12}	
$t = 6$							
Full space	12,072	1,164	18,964	45,929	3,000	–	1
Two step					256.01		0.95
Master	12,072	–	20,128	45,929	251.50	1.56×10^{12}	
Slave	12,072	24	20,103	45,879	4.51	1.57×10^{12}	
$t = 7$							
Full space	14,084	1,358	22,116	53,567	3,000	–	1
Two step					434.74		0.83
Master	14,084	–	23,474	53,567	431.30	1.68×10^{12}	
Slave	14,084	28	23,445	53,517	3.44	1.69×10^{12}	
$t = 8$							
Full space	16,096	1,552	25,268	61,205	3,000	–	1
Two step					669.08		0.59
Master	16,096	–	26,820	61,205	657.86	1.80×10^{12}	
Slave	16,096	32	26,787	61,155	11.22	1.80×10^{12}	

[a] Objective function values obtained by solving the full-space problem with 1% optimality gap and the sub-problems of the sequential approach with 0% optimality gap

[b] Solution quality means here the optimality gap (%) of the solution provided by CPLEX (full-space problem) or the gap between the master and slave problems in the sequential approach

Table 2.8: Computational results for minimum expected cost ($t \leq 4$).

	Variables						
	Binary	Integer	Continuous	Equations	Time (s)	Cost (€)[a]	Quality (%)[b]
t = 1							
Full space	2,012	194	3,204	7,739	0.27	9.30×10^{11}ă	1
Two step					0.25		0.20
Master	2,012	–	3,398	7,739	0.16	9.28×10^{11}	
Slave	2,012	4	3,393	7,689	0.09	9.30×10^{11}	
t = 2							
Full space	4,024	388	6,356	15,377	0.91	1.05×10^{12}	1
Two step					0.73		0.21
Master	4,024	–	6,744	15,377	0.53	1.05×10^{12}	
Slave	4,024	8	6,735	15,327	0.20	1.05×10^{12}	
t = 3							
Full space	6,036	582	9,508	23,015	18.27	1.17×10^{12}	1
Two step					3.11		0.36
Master	6,036	–	10,090	23,015	2.36	1.17×10^{12}	
Slave	6,036	12	10,077	22,965	0.75	1.17×10^{12}	
t = 4							
Full space	8,048	776	12,660	30,653	29.00	1.29×10^{12}	1
Two step					6.40		0.43
Master	8,048	–	13,436	30,653	5.95	1.29×10^{12}	
Slave	8,048	16	13,419	30,603	0.45	1.30×10^{12}	

[a] Objective function values obtained by solving the full-space problem with 1% optimality gap and the sub-problems of the sequential approach with 0% optimality gap

[b] Solution quality means here the optimality gap (%) of the solution provided by CPLEX (full-space problem) or the gap between the master and slave problems in the sequential approach

Table 2.9: Computational results for minimum expected cost ($t \geq 5$).

	Variables						
	Binary	Integer	Continuous	Equations	Time (s)	Cost (€)[a]	Quality (%)[b]
t = 5							
Full space	10,060	970	15,812	38,291	70.09	1.42×10^{12}	1
Two step					4.84	0.44	
Master	10,060	–	16,782	38,291	4.26	1.42×10^{12}	
Slave	10,060	20	16,761	38,241	0.58	1.42×10^{12}	
t = 6							
Full space	12,072	1,164	18,964	45,929	100.58	1.55×10^{12}	1
Two step					10.85		0.54
Master	12,072	–	20,128	45,929	8.57	1.55×10^{12}	
Slave	12,072	24	20,103	45,879	2.28	1.551×10^{12}	
t = 7							
Full space	14,084	1,358	22,116	53,567	235.34	1.67×10^{12}	1
Two step					11.07		0.54
Master	14,084	–	23,474	53,567	9.73	1.66×10^{12}	
Slave	14,084	28	23,445	53,517	1.34	1.67×10^{12}	
t = 8							
Full space	16,096	1,552	25,268	61,205	372.66	1.78×10^{12}	1
Two step					20.15		0.55
Master	16,096	–	26,820	61,205	15.31	1.78×10^{12}	
Slave	16,096	32	26,787	61,155	4.84	1.79×10^{12}	

[a] Objective function values obtained by solving the full-space problem with 1% optimality gap and the sub-problems of the sequential approach with 0% optimality gap

[b] Solution quality means here the optimality gap (%) of the solution provided by CPLEX (full-space problem) or the gap between the master and slave problems in the sequential approach

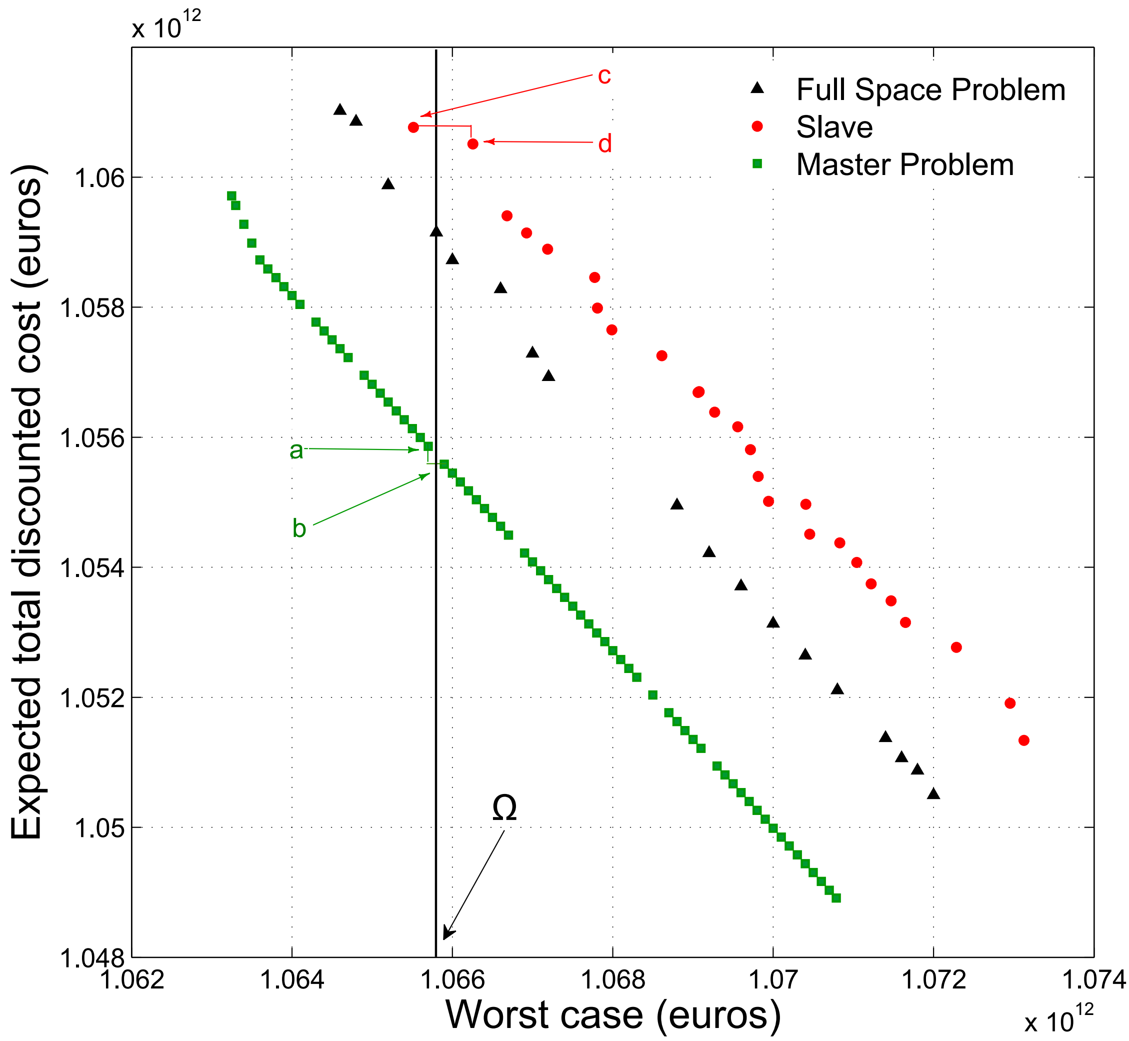


Figure 2.3: Pareto optimal solution curves for *t=2* using 2-step sequential approach and full space methods.

2.5.3.2 Pareto optimal solutions

Having proved the application and computational effectiveness of the proposed algorithm, we next use it to determine the Pareto set of the bi-criterion problem. Specifically, Figure 2.3 shows the results obtained with the full space method and the decomposition strategy for a two time period resolution. For a higher number of periods, the full-space method is only able to solve the extreme points of the curve, and as shown in the figure, there is a clear trade-off between the expected cost and financial risk. An improvement in one of the metrics can only be achieved by sacrificing the other. As observed, the solutions calculated by the master and slave problems provide valid lower and upper envelopes within which the Pareto set of (P) must fall.

As an illustrative example, consider the case in which a decision-maker wants to know the lower and upper bounds on the best possible expected cost for a worst case value equal to Ω. As shown in the figure, the minimum possible expected cost corresponds to that of solution b. This would be the cost attained if the Pareto set would be completely flat between a and b. This is indeed the best possible situation that could occur since it implies that no trade-off exists between both points. On the other hand, an upper bound on the cost is provided by the objective function value of solution c. This is because among all the feasible solutions of (P) with a worst case value below omega (*i.e.*, $WC < \Omega$) that have been calculated by the slave problem, c is the one with maximum expected cost. Note that the points above the upper bound envelope are non Pareto optimal, since they are dominated at least by those belonging to the upper bound set. On the other hand, solutions below the lower bound set are impossible to attain, since the master problem of our method provides a rigorous lower bound on the global solution of (P).

For the case of target values that fall below the minimum worst case value calculated by the slave problem, we can only provide lower bounds but not upper bounds to the problem. In fact, for that region of the search space we cannot even know if there exists a feasible solution. Thus, the minimum possible worst case that can be attained will always lie between the minimum worst case values provided by the master and slave problems. The main reason for that is that the master problem provides a rigorous lower bound, while the slave problem is a lissome bound more close in theory to the real value of the solution. In practice, the quality of the lower and upper bounds provided by the method will depend on the specific problem being solved. As observed, in our case these bounds are quit tight.

In Figure 2.4, the Pareto optimal solution curve corresponding to the slave upper bound problem for a six time period resolution of the stochastic MILP problem is presented. This represents a complex problem in which the full space method fails to provide the entire Pareto set. For the sake of simplicity, we only provide the upper bound curve in this case.

As in the previous case, there is a clear trade-off between expected cost and worst case . Each point of the Pareto set corresponds to a specific hydrogen network design. The minimum cost solution (point A) corresponds to the single objective problem that would arise from minimizing exclusively the expected total cost, as introduced in Section 2.4. The specific supply chain design of this solution is shown in Figure 2.5. This is also the solution that one would obtain with standard deterministic approaches. In this solution, the model suggests to produce liquid hydrogen *via* steam methane reforming and to construct plants in every potential location in order to reduce the material flows between the regions. This is mainly a consequence of the exceptional expected economic performance that steam methane reforming plants show nowadays in comparison with the rest of technologies, which is largely due to the competitive price of natural gas.

Nevertheless, the price of natural gas shows very high variability in contrast to its main economic competitor, the coal. Thus, as uncertainty is unveiled over the different

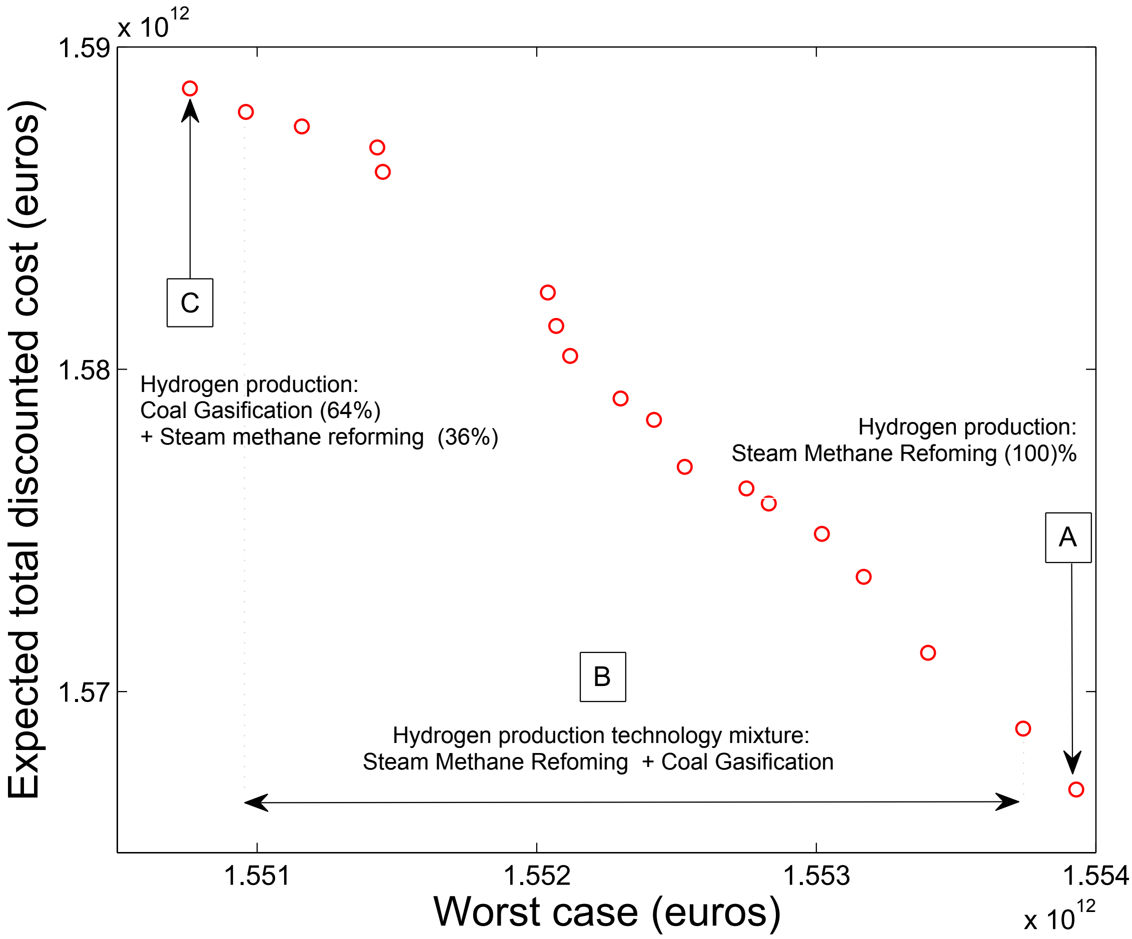


Figure 2.4: Pareto optimal solution curve for *t=6* using 2-step sequential approach. Slave problem solution.

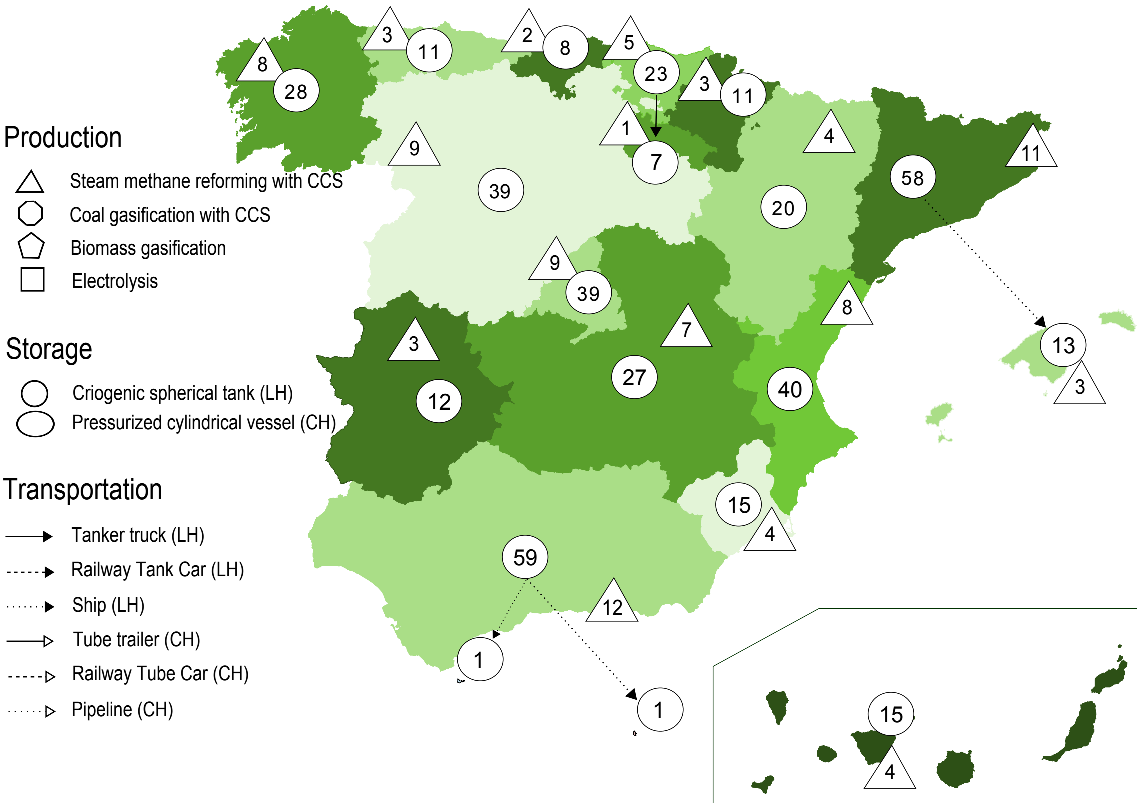


Figure 2.5: Hydrogen supply chain design for the minimum cost solution. The numbers inside the circles and triangles denote the number of facilities of each type.

scenarios represented, and by accounting for the worst case cost occurring in these realizations, the model resolves that the solution showing less variability over the time is the production *via* coal gasification. Therefore, the minimum worst case solution, which corresponds to point C of Figure 2.4 and to the hydrogen supply chain design depicted in Figure 2.6, entails a mixture of production technologies corresponding to a large extent to coal gasification for liquid hydrogen and some steam methane reforming plants also for liquid hydrogen. The points lying in between these extreme solutions correspond to specific designs involving different mixtures of steam methane reforming and coal gasification production technologies. Note that Figure 2.5 shows a more decentralized hydrogen network, while in Figure 2.6 a considerable increment on transportation links is observed. This is mainly due to the economic performance of steam methane reforming plants, which shows the economic advantages of building larger plants that implement this technology versus transporting the hydrogen.

Figure 2.7 depicts the cumulative probability curves of the feasible extreme solutions calculated by the slave problem. The aforementioned curves show for each possible target value imposed on the total cost, the probability of attaining a cost lower or equal to it. Note that according to the definition of financial risk, the cumulative probability corresponds to the difference between one and the risk associated with the given target

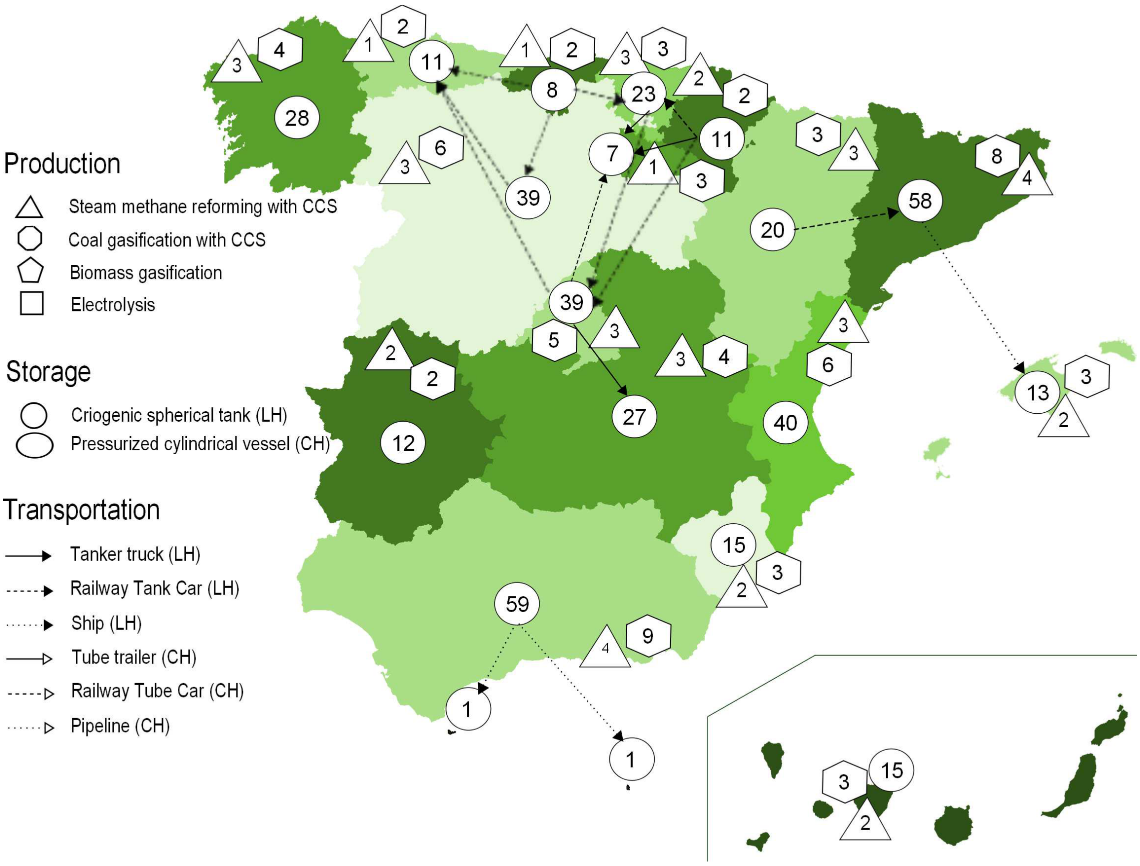


Figure 2.6: Hydrogen supply chain design for the minimum worst case solution. The numbers inside the circles and triangles denote the number of facilities of each type.

value, as it is depicted in Figure 2.7. The figure shows that for low cost levels (*i.e.*, lower than 1.557×10^{12} €) the minimum cost solution, which entails a hydrogen network formed by steam methane reforming plants, show a level of risk lower than the minimum worst case solution, which encompasses a mixture of production plants dominated by coal gasification ones. On the other hand, for high cost levels (*i.e.*, higher than 1.557×10^{12} €) the coal gasification technology shows less financial risk than steam methane reforming. Hence, both solutions represent different attitudes towards risk, being solution C the one with lower probability of unfavorable scenarios with high cost. Due to the inherent trade-off between expected economic performance and risk, there is no solution that performs better considering the whole range of target values. Therefore, the task of decision makers is to select among the Pareto points the one that better reflects their objectives.

It should be noted that the outcome of the analysis is indeed very sensitive to the specific data considered.

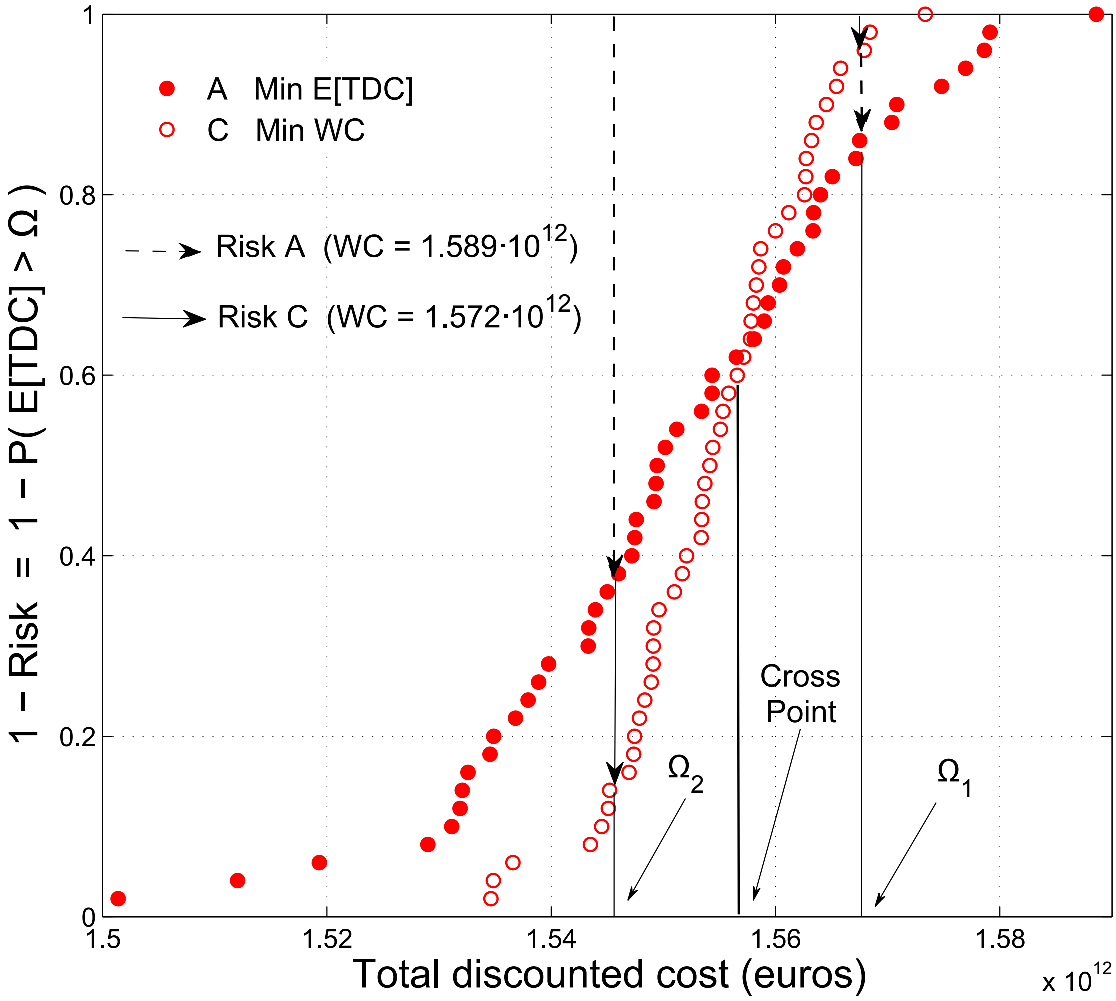


Figure 2.7: Cumulative risk curves associated with the extreme solutions of the slave problem for *t=6*.

2.6 Conclusions

This work has introduced a novel decision-support tool for managing the financial risk associated in the strategic design and planning of hydrogen supply chains under uncertainty in the operating costs *via* the worst case value. One of its additional novelties is the inclusion of all possible technologies available at the moment for hydrogen production, storage and transportation, into a single mathematical model. The framework proposed is able to evaluate all their possible combinations considering several periods of time and in the face uncertainties in the coefficients of the objective function using a multi-scenario systematic approach.

The problem has been mathematically posed as a multi-objective multi-scenario multi-period stochastic MILP formulation accounting for the minimization of the expected total discounted cost and the worst case value. Thus, the model explicitly incorporates the trade-off between risk and cost at the decision-making level.

Furthermore, a two-step sequential approach that exploits the specific mathematical structure of the model formulation has been introduced as a way to overcome the numerical difficulties associated with the application of the proposed strategy to large scale problems.

In Section 3.1 the importance of the infrastructure in determining the future of hydrogen as an energy vector has been presented. Also the existent optimization frameworks aimed at paving the way for this transition and important decision period have been reviewed. The approaches presented differ in the following characteristics:

- performance metrics considered (*i.e.*, the amount of objectives considered and their nature)
- geographical space resolution (*e.g.*, regional, country-level, *etc.*)
- time representation (*i.e.*, single-period or multi-period)
- uncertainty consideration (*i.e.*, deterministic or stochastic)

The mathematical model presented in Section 3.3 is a demand driven model, which considers capacity expansions and economies of scale for the production facilities, assumes that the raw material availability and its associated transport costs are negligible compared to the rest of the costs of the network and the uncertainty in the operating costs of the network reflecting energy market fluctuations.

As has been reported, stochastic frameworks optimizing the expected performance of problems where uncertainty is present in the coefficients of the objective function are equivalent to the optimization of the mean scenario. In that sense the functionalities of a more complex stochastic framework would not be useful unless its additional information is used. Our approach specifically handles this issue by appending to the objective function an additional criteria represented by a risk management metric with reported good computational performance and effectiveness.

Note that although the main goal of the model is to provide a robust framework for dealing with parametric uncertainty in the operating costs, the Pareto sets obtained as a result of the underlying formulation could be used for the methodological uncertainty analysis associated to single-objective economic optimization frameworks. This is here possible because, as mentioned earlier, the expected performance in these problems tends to be equivalent with the mean scenario, which is the solution of its deterministic problem counterpart.

On the computational side, the solution strategy detailed in Section 3.4 has been shown from numerical examples to be very efficient, providing solutions with an optimality gap lower than 1% in a fraction of the CPU time (*i.e.*, at least one order of magnitude faster) required by the original formulation. The strategy was applied to a case study consisting of:

- 19 potential locations
- 8 different production modes
- 2 types of storage facilities
- 6 transportation modes

The most complex instance of the problem solved for 8 time periods gives rise to 25,268 continuous, 16,096 binary and 1,552 integer variables, showing that almost half of the total number of variables of the problem are binary and integer. In this sense the decomposition method can also be applied to similar SC models presented so far in the literature, thus allowing for its practical implementation in problems of larger size.

Finally, the effectiveness of the proposed approach as a decision-making tool capable of providing insights into the SC design problem has been also highlighted. Results indicate that for a given level of cost, lower levels of risk can be obtained by switching from steam methane reforming to coal gasification production plants at the design step.

The financial risk optimization using worst case scenario as performance metric has the capability of informing decision-makers about the total cost value at which both (*i.e.*, mean and risk averse) supply chain designs are equivalent from a risk point of view. This is achieved through the representation of the cumulative probability distribution functions of the objectives optimized. These in turn, unveil the robustness advantages of the strategy dominated by coal gasification plants in the event of higher energy costs. The results also show the clear financial advantage that steam methane reforming plants present for low energy price scenarios. From a purely economic point of view, the production of liquid hydrogen appears to be superior in all instances if new infrastructure needs to be built.

It is also seen that the share of the variations considered in the operating costs have a share of 2 % in the total discounted cost of the network. If these variations were extended to other potentially affected parameters of the network the share would presumably increase. In any situation, the presented results illustrate the shift in the network designs that can be expected even in the event of small differences in the overall budget.

The insights obtained in the numerical analysis might change according to the input data. However, the method introduced is general enough to be adapted to any particular situation. The tool presented is intended to enhance our knowledge on how to establish optimal SC networks capable of dealing with uncertainties in realistic problems.

Notation

Indices

e	scenarios
i	hydrogen form
g	potential locations
l	transportation mode
p	manufacturing technologies
s	storage technologies
t	time period

Sets

$IL(l)$	set of hydrogen forms that can be transported *via* transportation mode l
$IP(p)$	set of hydrogen forms that can be produced *via* technology p
$IS(s)$	set of hydrogen forms that can be stored *via* technology s
$LI(i)$	set of transportation modes for hydrogen form i
$PI(i)$	set of technologies that can produce hydrogen form i
$SI(i)$	set of storage technologies for hydrogen form i
$SGG(gg')$	set of allowable maritime links
$SGG'(gg')$	subset of maritime links that cannot be covered by road transportation units

Parameters

av_l	availability of transportation mode l
cc_{lt}	capital cost of transport mode l in period t
cud_{lte}	maintenance cost of transportation mode l in period t per unit of distance traveled in scenario e
$\overline{D_{gt}}$	total demand of hydrogen in location g in period t
$distance_{gg'}$	average distance traveled between locations g and g'
$dsat$	demand satisfaction level to be fulfilled
$fuelc_l$	fuel consumption of transportation mode l
$fuelp_{lte}$	pfuel price for transportation mode l in period t in scenario e
ge_{lte}	general expenses of transportation mode l in period t
ir	interest rate
$lutime_l$	loading/unloading time of transportation mode e
$\overline{PC_p^{PL}}$	upper bound on the capacity expansion of manufacturing technology p
$\underline{PC_p^{PL}}$	lower bound on the capacity expansion of manufacturing technology p
θ	average storage period
τ	minimum desired percentage of capacity to be used
$prob_e$	probability of occurrence of scenario e

$\overline{QC_{gg'l}}$	upper bound on the flow of materials between locations g and g' *via* transportation model l
$\underline{QC_{gg'l}}$	lower bound on the flow of materials between locations g and g' *via* transportation model l
$\overline{SC_{s}^{ST}}$	upper bound on capacity expansion of storage technology s
$\underline{SC_{s}^{ST}}$	lower bound on capacity expansion of storage technology s
$speed_l$	average speed of transportation mode l
$tcap_l$	capacity of transport mode l
upc_{igpte}	value of unit production cost of hydrogen form i produced *via* technology p in location g in period t for scenario e
usc_{igste}	unit storage cost of hydrogen form i stored *via* technology s in location g in period t in scenario e
$upoc_{te}$	unit operating cost of the pipelines in scenario e
$usoc_{te}$	unit operating costs for maritime transportation in scenario e
SOC_{te}	operating cost of ships in period t and scenario e
ROC_{te}	operating costs associated with road transportation technologies and railway in period t and scenario e
$wage_{lte}$	driver wage of transportation mode l in period t in scenario e
α_{gpt}^{PL}	fixed investment term associated with manufacturing technology p installed in location g in period t
α_{gst}^{ST}	fixed investment term associated with storage technology s installed in location g in period t
β_{gpt}^{PL}	variable investment term associated with manufacturing technology p installed in location g in period t
β_{gst}^{ST}	variable investment term associated with storage technology s installed in location g in period t

Variables

C_{gpt}^{PL}	capacity of manufacturing technology p in location g in period t
C_{gst}^{ST}	capacity of storage technology s in location g in period t
CE_{gpt}^{PL}	capacity expansion of production technology p in location g in period t
CE_{gst}^{ST}	capacity expansion of storage technology s in location g in period t
D_{igt}	amount of hydrogen form i distributed in location g in period t
FC_{te}	fuel cost in period t in scenario e
FCC_t	facility capital cost in period t
FOC_{te}	facility operating cost in period t in scenario e
GC_{te}	general cost in period t in scenario e
LC_{te}	labor cost in period t in scenario e
MC_{te}	maintenance cost in period t in scenario e
$upcc_t$	unit transportation cost of pipelines in period t
PCC_t	pipeline capital cost in period t and scenario e
POC_{te}	pipeline operating cost in period t and scenario e

SOC_{te}	ship operating cost in period t and scenario e
PR_{igpt}	production of hydrogen mode i *via* technology p in period t in location g
$Q_{igg'lt}$	flow of hydrogen mode i *via* transportation mode l between locations g and g' in period t
S_{igst}	amount of hydrogen in physical form i stored *via* technology s in location g in period t
TC_{te}	total amount of money spent in period t and scenario e
TCC_t	total transportation capital cost in period t
TDC_e	total discounted cost and scenario e
TOC_{te}	transportation operating cost in period t in scenario e
$E[TDC]$	Expected total discounted cost of the network
WC	Worst case cost of the network

Integer variables

N^{PL}_{gpt}	number of plants of type p installed in location g in period t (integer variable)
N^{ST}_{gst}	number of storage facilities of type s installed in location g in period t (integer variable)
N^{TR}_{lt}	number of transportation units of type l purchased in period t (integer variable)

Binary variables

$X_{gg'lt}$	binary variable (1 if a link between locations g and g' using transportation technology l is established, 0 otherwise)

CHAPTER 3

MULTI-OBJECTIVE OPTIMIZATION OF A HYDROGEN SUPPLY CHAIN – DETERMINISTIC APPROACH FOR OBJECTIVE REDUCTION

If you are Noah, and your ark is about to sink, look for the elephants first, because you can throw over a bunch of cats, dogs, squirrels, and everything else that is just a small animal and your ark will keep sinking. But if you can find one elephant to get overboard, you are in much better shape

Vilfredo Pareto, 1848 – 1923

3.1 Introduction

Assessing the environmental performance of hydrogen infrastructures is essential for determining their practical viability. Previous optimization approaches for hydrogen networks, like the one presented in the previous chapter, have mostly focused on optimizing the economic performance. Meanwhile, the few works in the literature attending environmental impacts relied on a single metric to measure the environmental performance. Given that the study of environmental impact indicators is an active area of research and constitutes a major subject under current debate, this approach can be inadequate as it may leave relevant environmental criteria out of the analysis. We propose herein a novel framework for optimizing hydrogen supply chains (SC) according to several environmental indicators. The model builds on the formulation of the problem presented in Chapter 2, and by bringing the problem to its deterministic form, we are then able to append the building blocks of the environmental performance analysis. Our method comprises two steps. In step one, we formulate a multi-objective mixed-integer linear program (MILP) that accounts for the simultaneous minimization of the most relevant life cycle assessment (LCA) impacts. Principal Component Analysis (PCA) is next employed in the post-optimal analysis of the MILP in order to facilitate the interpretation and analysis of its solution space. We demonstrate the capabilities of this approach through its application to the design of the future (potential) hydrogen SC in Spain. Note that the present work was also published as an original research article (Sabio et al., 2012).

3.1.1 Infrastructure optimization and the future of a hydrogen economy

A major obstacle in the transition towards a hydrogen economy is the lack of infrastructures to produce, store and deliver hydrogen. Even though hydrogen is an abundant chemical element, it does not exist in a free form. Therefore, it must be generated *via* different technologies, such as steam methane reforming, coal gasification, water electrolysis, and biomass gasification, among others. These technologies differ in capital investments, required feedstocks, production cost (Balat and Kirtay, 2010) and environmental performance (Koroneos et al., 2004, 2005; Spath and Mann, 2004, 2001). Designing efficient hydrogen infrastructures requires the simultaneous assessment of all these alternatives and subsequent identification of the best combination of them in terms of some pre-defined criteria. Mathematical tools can assist decision-makers in this task by automating the search for optimal solutions, thereby guiding policy makers towards the adoption of hydrogen infrastructures with improved economic and environmental performance.

Most of the optimization approaches for hydrogen networks available in the literature seek to optimize the economic performance. Almansoori and Shah (2006) developed a deterministic mathematical programming model to minimize the total daily cost of the future British hydrogen SC. Ingason et al. (2008) presented an optimization model for the design of a hydrogen network in Iceland that minimized the annual cost of the investments. The model of Lin et al. (2008a) identified the configuration of a SC producing the least-cost hydrogen for South California. Kim et al. (2008) proposed a stochastic formulation to examine the total daily cost of the South Korean hydrogen SC under demand uncertainty. Sabio et al. (2010) developed a stochastic model to optimize hydrogen SCs under uncertainty in the operating costs.

Optimizing exclusively the economic performance may lead to solutions that do not fully exploit the environmental benefits of moving towards a hydrogen-based energy system (Hugo et al., 2005). It is therefore nowadays becoming increasingly clear that environmental concerns must be accounted for along with economic criteria in the optimization of hydrogen infrastructures. Multi-criteria decision-making, and particularly multi-objective optimization, provide a systematic framework to accomplish this task.

3.1.2 Multi-objective optimization (MOO) coupled with principal components analysis (PCA)

Hugo et al. (2005) proposed a MILP model for the long-term strategic planning of a multi-echelon hydrogen network that optimizes both economic and environmental criteria (*i.e.*, greenhouse gas (GHG) emissions). The bi-criterion model of Guillén-Gosálbez et al. (2010) takes into account the minimization of the damage to human health caused by climate change and the total daily cost of hydrogen SC. Li et al. (2008) developed a MILP model for optimizing the future hydrogen infrastructure in China that simultaneously minimizes the associated GHG emission and maximizes the profit of the hydrogen SC.

The aforementioned works restrict the analysis to two objectives: one economic indicator and one environmental metric. This consideration can lead to solutions in which a single environmental damage (*e.g.*, global warming potential) is reduced at the expense of increasing others (*e.g.*, acidification, respiratory effects, *etc.*). A more comprehensive and holistic environmental assessment of hydrogen SCs requires the simultaneous consideration of several damages in the decision-making process.

Unfortunately, increasing the number of objectives leads to problems whose solutions are difficult to visualize and analyze. Objective reduction techniques (Deb and Saxena, 2005; Brockhoff and Zitzler, 2009; Guillén-Gosálbez, 2011b) attempt to overcome this limitation by identifying redundant metrics that can be omitted while still preserving the mathematical structure of the problem to the extent possible. Further, these methods allow uncovering relationships between environmental impacts, enhancing our understanding on the environmental performance of the hydrogen infrastructure.

In this work, we integrate multi-objective optimization with PCA (Jackson, 2003) to address the environmentally conscious design of hydrogen networks. Multi-objective optimization enables the systematic calculation of optimal trade-off solutions, whereas PCA identifies redundant objectives that can be left out of the analysis, shedding light on the interactions between the environmental damages caused by the hydrogen infrastructure.

The capabilities of our approach are illustrated through a case study that addresses the design of the future hydrogen SC to be established in Spain. The model includes three main hydrogen production technologies (*i.e.*, steam methane reforming, coal gasification, and water electrolysis), two types of hydrogen storing (*i.e.*, cryogenic spherical tanks for liquid H_2 and pressurized vessels for compressed H_2), and six transportation alternatives (*i.e.*, tanker trucks, railway tank cars, ships, tube trailers, railway tube cars, and pipelines). The environmental performance of the system is evaluated according to a set of life cycle impacts measured following a cradle-to-gate approach. Numerical results show that several environmental effects of the hydrogen network are highly correlated, which makes it possible to focus our attention on a reduced set of damages. Our approach identifies those technological alternatives representing the optimal trade-off between the economic and environmental performance of the hydrogen network.

The remainder of this chapter is organized as follows. In Section 3.2, the problem under study is formally stated, and the assumptions made are briefly presented. In Section 3.3, we describe the equations of the multi-objective MILP derived to tackle the SC design problem. In Section 3.4, we present a combined method that integrates multi-objective optimization and PCA. In Section 3.5, the proposed approach is applied to a real case study based on the future Spanish hydrogen SC, for which valuable insights are obtained. The conclusions of the work are finally drawn in Section 3.6 of this chapter.

3.2 Problem statement

The goal of the design problem addressed in this chapter is to determine the optimal configuration of a three-echelon hydrogen SC for vehicle use (production-storage-market) in terms of cost and damage to the environment. For our analysis, we will consider a generic hydrogen SC superstructure like the one depicted in Figure 3.1. This network comprises a set of production plants (pentagons), and a set of storage facilities (circles), where hydrogen is stored before being delivered to end customers (rectangles).

A region of interest divided into a set of potential locations is considered. These potential locations correspond to different sub-regions in which a given hydrogen demand must be satisfied. We consider that the set of potential locations along with the associated geographical distribution of the demand are input data to the problem.

The design problem can be formally stated as follows. Given are the hydrogen demand, a fixed time horizon and set of time periods, a set of available production, storage and transportation technologies, the capacity limitations of plants and storage facilities, the costs associated with the network operation (production, transportation and inventory costs), the investment cost, and the interest rate. The goal is to determine the following decisions:

- the number, type, location and capacity of plants and storage facilities, along with the number and type of transportation units (*e.g.*, tanker trucks, railway tube cars,

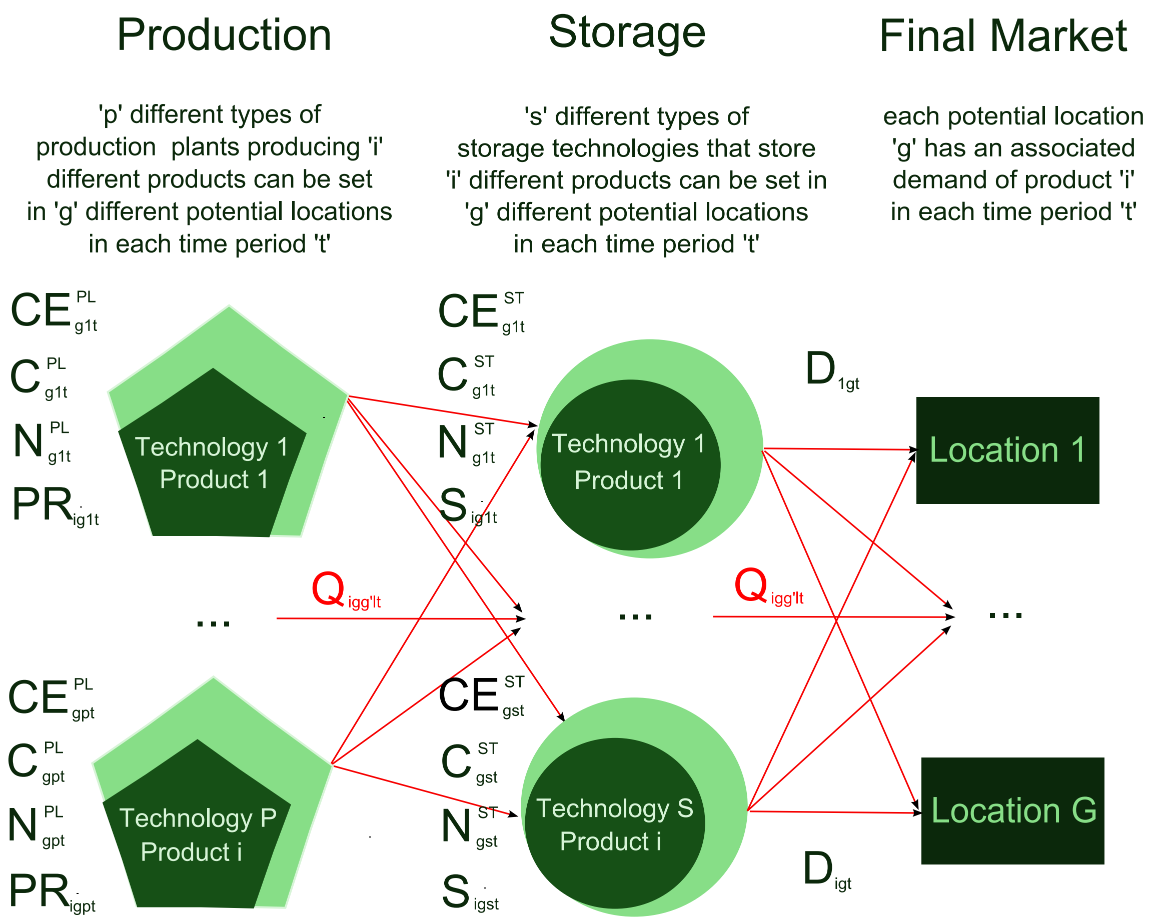


Figure 3.1: Structure of three echelon supply chain

etc.) and transportation links to be established between the potential locations;

- the production rates at the plants, inventory levels at the storage facilities and flows of hydrogen between plants and storage facilities;

so as to simultaneously minimize the total cost and the associated damage to the environment.

3.3 Mathematical model

The MILP formulation derived to tackle the problem presented above is based on that introduced by Almansoori and Shah (2006) and Guillén-Gosálbez et al. (2010). The mathematical model considers all possible configurations of the future hydrogen SC as well as all technological aspects associated with the SC performance, including several production and storage technologies and transportation modes. For the sake of completeness of this work, we provide next an outline of the mathematical formulation. Further details can be found in the works by Guillén-Gosálbez et al. (2010) and Sabio et al. (2010). Particularly, we will first discuss some general features of the model before immersion into a detailed description of its equations.

Production plants

Hydrogen can be produced in different ways. In this chapter, we consider only three technologies for hydrogen production plants: steam methane reforming (SMR), coal gasification (CG) and water electrolysis (WE). Note, however, that the model could be easily extended in order to account for more technologies (*e.g.*, biomass gasification, biomass pyrolysis, photobiological H_2 production, nuclear-powered electrolysis *etc.*). Steam methane reforming of natural gas is the leading technology now and the pathway by which most hydrogen is made today. Konieczny et al. (2008) reported that nearly 48% of hydrogen is produced *via* SMR from natural gas. Nowadays, SMR is regarded as the most profitable way to produce hydrogen (Balat and Kirtay, 2010). However, during the conversion of natural gas into hydrogen, a SMR plant emits large amounts of GHGs mostly CO_2 and CH_4 (Spath and Mann, 2001). The production cost of hydrogen *via* coal gasification is about twice as much as from natural gas. Coal is much cheaper than natural gas, but the capital and operating costs of coal gasification are higher than those associated with a SMR plant (Padró and Putsche, 1999). The most significant environmental impact associated with coal gasification is given by the emissions of CO_2 and SO_2 (Koroneos et al., 2005). Electrolysis of water is a promising technology to produce hydrogen, since it can be coupled with a renewable source of energy (*e.g.*, wind power), thereby avoiding the usage of fossil fuels. In this work, we consider that the electrolyzer employs electricity from wind turbines. This system is more expensive than SMR and coal gasification, but offers significant environmental benefits.

Storage facilities

The proposed model considers two possible alternatives for physical storage of hydrogen: compressed hydrogen storage and liquefied hydrogen storage. Liquefied gas requires less volume than compressed hydrogen. However, the liquefaction process consumes more utilities and demands additional equipment. The expenditures associated with the liquefaction and compression of hydrogen are included in the unit production and capital cost of the production facilities.

Transportation technologies

The terrestrial transport of compressed and liquid hydrogen can be carried out using trucks or railroad cars. In addition, pipelines can be constructed to deliver compressed hydrogen. Ships can also be freighted in order to supply maritime regions with hydrogen. These ships can transport hydrogen in both physical forms: either compressed gas or liquid.

3.3.1 Mass balance constraints

The mass balance must be satisfied for every hydrogen form i, liquid or gas, in each potential location g and time period t. Thus, for every location, the sum of the initial inventory $S_{i,g,s,t-1}$ plus the amount produced ($PR_{i,g,p,t}$) and the input flow rate ($Q_{i,g',g,l,t}$) must equal the final inventory ($S_{i,g,s,t}$) plus the amount delivered to the customers ($D_{i,g,t}$) and the output flow rate ($Q_{i,g,g',l,t}$) of hydrogen:

$$\begin{aligned}&\sum_{s\in SI(i)} S_{i,g,s,t-1} + \sum_{p\in PI(i)} PR_{i,g,p,t} + \sum_{g'\neq g}\sum_{l\in LI(i)} Q_{i,g',g,l,t}\\ &= \sum_{s\in SI(i)} S_{i,g,s,t} + D_{igt} + \sum_{g'\neq g}\sum_{l\in LI(i)} Q_{i,g,g',l,t} \qquad \forall i,g,t\end{aligned} \tag{1}$$

In Eq. (1), $SI(i)$ represents the subset of storage technologies that can be used for product form i, $LI(i)$ is the subset of transport modes that can transport product form i, and $PI(i)$ denotes the production facilities that can produce product form i.

3.3.2 Capacity constraints

3.3.2.1 Plants

In Eq. (2), the total production rate of hydrogen form i in location g produced *via* technology p in period t ($PR_{i,g,p,t}$) must be lower than the existing capacity of the plant and higher than a minimum desired percentage, τ, of the capacity installed. The capacity of each production technology p at location g in period t is represented by $C^{PL}_{g,p,t}$.

$$\tau\ C^{PL}_{g,p,t} \leq \sum_{i \in IP(p)} PR_{i,g,p,t} \leq C^{PL}_{g,p,t} \qquad \forall g, p, t \tag{2}$$

In this equation, $IP(p)$ denotes the subset of hydrogen forms that can be produced using technology p. The capacity of each technology p in period t is calculated by adding the expansion in capacity ($CE^{PL}_{g,p,t}$) executed in period t and location g to the existing capacity at the end of the previous period:

$$C^{PL}_{g,p,t} = C^{PL}_{g,p,t-1} + CE^{PL}_{g,p,t} \qquad \forall g, p, t \tag{3}$$

Equation (4) is applied to limit capacity expansions within lower and upper bounds. These bounds are obtained multiplying the number of plants installed (integer variable $N^{PL}_{g,p,t}$) with the minimum and maximum capacities of each technology p ($\underline{PC^{PL}_p}$ and $\overline{PC^{PL}_p}$, respectively).

$$\underline{PC^{PL}_p} N^{PL}_{g,p,t} \leq CE^{PL}_{g,p,t} \leq \overline{PC^{PL}_p} N^{PL}_{g,p,t} \qquad \forall g, p, t \tag{4}$$

3.3.2.2 Storage facilities

Equation (5) forces the total inventory of product in form i kept at the end of period t in the storage facilities of type s installed in location g ($S_{i,g,s,t}$), to be lower than the available capacity. Here, $C^{ST}_{g,s,t}$ represents the storage capacity of product form i during period t in location g associated with storage technology s.

$$\sum_{i \in IS(s)} S_{i,g,s,t} \leq C^{ST}_{g,s,t} \qquad \forall g, s, t \tag{5}$$

This equation constrains the amount of hydrogen delivered from the storage facility to the customers to be lower than its capacity. In this equation, $IS(s)$ represents the subset of product forms i that can be stored by technology s. Our model makes use of the storage period θ, as previously introduced by Almansoori and Shah (2006), which is multiplied by two in order to cover fluctuations in both supply and demand as well as plant interruptions (Simchi-Levi et al., 2000):

$$2\ (\theta\ D_{igt}) \leq \sum_{s \in SI(i)} C^{ST}_{g,s,t} \qquad \forall i, g, t \tag{6}$$

In Eq. (6), $SI(i)$ denotes the subset of storage technologies s that can handle product forms i. Similarly as with the manufacturing plants, the capacity of storage technology s at any time period t is determined from the previous one and the expansion in capacity executed in the same period:

$$C^{ST}_{g,s,t} = C^{ST}_{g,s,t-1} + CE^{ST}_{g,s,t} \qquad \forall g, s, t \tag{7}$$

Finally, the capacity expansion of the storage facilities $CE_{g,s,t}^{ST}$ is bounded within lower and upper limits.

$$\underline{SC_s^{ST}} N_{g,s,t}^{ST} \leq CE_{g,s,t}^{ST} \leq \overline{SC_s^{ST}} N_{g,s,t}^{ST} \qquad \forall g, s, t \tag{8}$$

3.3.3 Transportation constraints

To define these constraints, we introduce the binary variable $Y_{g,g',l,t}$ that represents the existence or absence of a transportation link of type l (*e.g.*, tanker trucks, railway tube cars, *etc.*) between locations g and g' in time period t.

$$\begin{aligned} \underline{QC_{g,g',l}} Y_{g,g',l,t} &\leq \sum_{i \in IL(l)} Q_{i,g,g',l,t} \\ &\leq \overline{QC_{g,g',l}} Y_{g,g',l,t} \qquad \forall g, g'(g \neq g'), l \neq \text{pipeline}, t \end{aligned} \tag{9}$$

When this binary variable is zero, there is no flow of materials *via* transportation technology l between g and g'. On the other hand, when the binary variable is one, it is possible to transport materials within some lower $\underline{QC_{g,g',l}}$ and upper bounds $\overline{QC_{g,g',l}}$. In Eq. (9), $IL(l)$ denotes the subset of product forms i that can be transported by transport mode l.

Equation (10) is similar to Eq. (9), but applies only to pipelines. Specifically, we assume that if a pipeline is constructed, then the associated transportation link will remain open over the entire time horizon:

$$\begin{aligned} \sum_{t'<t+1} \underline{QC_{g,g',l}} Y_{g,g',l,t'} &\leq \sum_{i \in IL(l)} (Q_{i,g,g',l,t} + Q_{i,g',g,l,t}) \\ &\leq \sum_{t'<t+1} \overline{QC_{g,g',l}} Y_{g,g',l,t'} \qquad \forall g, g'(g \neq g'), l = \text{pipeline}, t \end{aligned} \tag{10}$$

Furthermore, we cannot establish more than one transportation link using pipelines between two given regions during the entire time horizon:

$$\sum_{t'<t+1} Y_{g,g',l,t'} \leq 1 \qquad \forall g, g'(g \neq g'), l = \text{pipeline}, t \tag{11}$$

We assume that a region can either import or export hydrogen, but not both at the same time. This is because if a region cannot satisfy its needs with internal production, it will not export to other locations:

$$Y_{g,g',l,t} + Y_{g',g,l,t} \leq 1 \qquad \forall g, g'(g \neq g'), l, t \tag{12}$$

Eqs. (13) and (14) are added to handle maritime transportation. These equations force binary variable $Y_{g,g',l,t}$ (denoting the existence of transportation links) to take a zero value in some specific cases in order to prevent ships from transporting materials between locations with no harbors. These constraints avoid also road transportation between harbors that are not connected by roads.

$$\begin{aligned} & Y_{g,g',l,t} = 0 \qquad \forall l, g, g' \in LG \\ & LG = \{l, g, g' : (l = \text{ship}) \wedge ((g, g') \notin SGG(g, g'))\} \end{aligned} \tag{13}$$

$$\begin{aligned} & Y_{g,g',l,t} = 0 \qquad \forall l, g, g' \in LG' \\ & LG' = \{l, g, g' : (l \neq \text{ship,pipeline}) \wedge ((g, g') \in SGG'(g, g'))\} \end{aligned} \tag{14}$$

In these constraints, $SGG(g, g')$ is the subset of coastal regions with operating harbors, whereas $SGG'(g, g')$ is the subset of coastal regions with operating harbors (*i.e.*, $SGG'(g, g') \leq SGG(g, g')$) that cannot be connected through road transportation units.

3.3.4 Demand satisfaction constraint

The total amount of hydrogen consumed ($D_{i,g,t}$) is constrained to be lower than the total hydrogen demand ($\overline{D_{g,t}}$) in each location and period, and higher than a given minimum demand satisfaction level ($dsat$) :

$$\overline{D_{g,t}}\ dsat \leq \sum_i D_{i,g,t} \leq \overline{D_{g,t}} \qquad \forall g, t \tag{15}$$

3.3.5 Objective function equations

The model seeks to optimize the economic and environmental performance of the hydrogen SC. The economic objective is represented by the total discounted cost (TDC), whereas the environmental impact is quantified according to the LCA principles.

3.3.5.1 Total discounted cost

For the sake of simplicity, the economic performance of the SC is quantified according to the total discounted cost. Note, however, that more sophisticated financial metrics could be easily incorporated into the SC model, as was done in the past by the authors (Guillén-Gosálbez et al., 2007; Laínez et al., 2007). The total discounted cost is calculated as the summation of the discounted costs associated with each time period t:

$$TDC = \sum_t \frac{TC_t}{(1 + ir)^{t-1}} \tag{16}$$

In Eq. (16), ir represents the interest rate and TC_t is the total amount of money spent in period t , which includes the capital (FCC_t, TCC_t) as well as operating costs (FOC_t,

TOC_t) given by the production, storage and transportation facilities of the network:

$$TC_t = FCC_t + FOC_t + TCC_t + TOC_t \qquad \forall t \tag{17}$$

3.3.5.2 Facility capital cost

The facility capital cost in period t (FCC_t) is determined from the capacity expansions made in the manufacturing plants and storage facilities during that period:

$$\begin{aligned} FCC_t = & \sum_g \sum_p \left(\alpha^{PL}_{g,p,t} N^{PL}_{g,p,t} + \beta^{PL}_{g,p,t} CE^{PL}_{g,p,t} \right) \\ & + \sum_g \sum_s \left(\alpha^{ST}_{g,s,t} N^{ST}_{g,s,t} + \beta^{ST}_{g,s,t} CE^{ST}_{g,s,t} \right) \qquad \forall t \end{aligned} \tag{18}$$

Here, the parameters, $\alpha^{PL}_{g,p,t}$, $\beta^{PL}_{g,p,t}$, $\alpha^{ST}_{g,s,t}$ and $\beta^{ST}_{g,s,t}$ are the fixed and variable investment terms corresponding to plants and storage facilities, respectively.

3.3.5.3 Facility operating cost

The facility operating cost term is obtained by multiplying the unit production and storage costs ($upc_{i,g,p,t}$ and $usc_{i,g,s,t}$, respectively) with the corresponding production rates and average inventory levels:

$$\begin{aligned} FOC_t = & \sum_i \sum_g \sum_{p \in PI(i)} upc_{i,g,p,t} PR_{i,g,p,t} \\ & + \sum_i \sum_g \sum_{s \in SI(i)} usc_{i,g,s,t} \left(\theta \ D_{i,g,t} \right) \qquad \forall t \end{aligned} \tag{19}$$

3.3.5.4 Transportation capital cost

The transportation capital cost, which includes the cost of the trucks and railcars is calculated *via* Eq. (20):

$$TCC_t = \sum_{l \neq \text{ship,pipeline}} N^{TR}_{l,t} \, cc_{l,t} + PCC_t \qquad \forall t \tag{20}$$

Here, PCC_t is the pipeline capital costs, $cc_{l,t}$ represents the capital cost associated with transport mode l in period t, and $N^{TR}_{l,t}$ is an integer variable denoting the total number of transportation units of type l purchased in period t. Note that ships and pipelines are excluded from the first term of the summation. This is because the model assumes that ships are hired for carrying out the specific transportation tasks (*i.e.*, outsourcing).

The capital cost of pipelines is calculated *via* the following equation:

$$PCC_t = \sum_{g} \sum_{g' \neq g} \sum_{l=\text{pipeline}} upcc_t \; Y_{g,g',l,t} \; distance_{g,g'} \qquad \forall t \tag{21}$$

where $upcc_t$ is the unit capital cost of the pipeline per unit of length built and $distance_{g,g'}$ denotes the distance between potential locations g and g'.

The average number of trucks and/or railcars required to satisfy a certain flow between different locations is computed from the flow rate of products between the locations ($Q_{i,g,g',l,t}$), the transportation mode availability (av_l), the capacity of a transport container ($tcap_l$), the average distance ($distance_{g,g'}$), the average speed ($speed_l$) and the loading/unloading time ($lutime_l$), as stated in Eq. (22):

$$\begin{aligned} \sum_{t'<t+1} N^{TR}_{l,t'} \geq \sum_{i \in IL(l)} \sum_{g} \sum_{g' \neq g} \sum_{t} \frac{Q_{i,g,g',l,t}}{av_l \; tcap_l} \\ \left(\frac{2 \; distance_{g,g'}}{speed_l} + lutime_l \right) \qquad \forall l \neq \text{ship,pipeline} \end{aligned} \tag{22}$$

The total number of transportation units available in any period t includes the ones purchased in the same period t as well as those acquired in the past (*i.e.*, in previous periods t'). Therefore, the left hand side of the inequality in Eq. (22) represents the summation of all the transportation units purchased in all the time periods t' up to the actual period t (*i.e.*, $t' = t$). Also here, $IL(l)$ denotes the subset of product forms i that can be transported by transport mode l. For simplicity, this work assumes that each transportation facility can only operate between two pre-defined locations. For this reason the distance between locations g and g' ($distance_{g,g'}$) is multiplied by two in order to account for the return journey of the trucks/ railcars.

3.3.5.5 Transportation operating cost

The total operating cost associated with the transportation tasks carried out in period t (TOC_t) is determined from Eq. (23):

$$TOC_t = ROC_t + POC_t + SOC_t \qquad \forall t \tag{23}$$

where ROC_t, POC_t and SOC_t are the operating costs associated with road transportation technologies and railway, pipelines and ships, respectively. The first term includes the fuel (FC_t), labor (LC_t), maintenance (MC_t) and general costs (GC_t):

$$ROC_t = FC_t + LC_t + MC_t + GC_t \qquad \forall t \tag{24}$$

The fuel cost is a function of the fuel price ($fuelp_{l,t}$) and fuel consumption:

$$FC_t = \sum_g \sum_{g' \neq g} \sum_{l \neq \text{ship, pipeline}} \sum_{i \in IL(l)} fuelp_{l,t} \frac{2\ distance_{g,g'}\ Q_{i,g,g',l,t}}{fuelc_l\ tcap_l} \qquad \forall t \tag{25}$$

In Eq. (25), the fractional term represents the fuel usage, and it is determined from the total distance traveled in a trip ($2\ distance_{g,g'}$), the fuel consumption of transport mode l ($fuelc_l$) and the number of trips made per period of time ($\frac{Q_{i,g,g',l,t}}{tcap_l}$). The labor transportation cost is a function of the driver wage ($wage_{l,t}$) and total delivery time (*i.e.*, the term inside the brackets):

$$\begin{aligned} LC_t = & \sum_g \sum_{g' \neq g} \sum_{l \neq \text{ship, pipeline}} \sum_{i \in IL(l)} wage_{l,t} \\ & \times \left[\frac{Q_{i,g,g',l,t}}{tcap_l} \left(\frac{2\ distance_{g,g'}}{speed_l} + lutime_l \right) \right] \qquad \forall t \end{aligned} \tag{26}$$

The maintenance cost, which accounts for the general maintenance of the transportation systems, is a function of the cost per unit of distance traveled ($cud_{l,t}$) and total distance driven:

$$MC_t = \sum_g \sum_{g' \neq g} \sum_{l \neq \text{ship, pipeline}} \sum_{i \in IL(l)} cud_{l,t} \frac{2\ distance_{g,g'}\ Q_{i,g,g',l,t}}{tcap_l} \qquad \forall t \tag{27}$$

The general cost includes the transportation insurance, license and registration, and outstanding finances. It can be determined from the unit general expenses ($ge_{l,t}$) and number of transportation units as follows:

$$GC_t = \sum_{l \neq \text{ship, pipeline}} \sum_{t' \leq t} ge_{l,t} N^{TR}_{l,t'} \qquad \forall t \tag{28}$$

Equation (29) determines the pipeline operating costs from the unit operating cost of the pipelines($upoc_t$) and freight to be delivered.

$$POC_t = \sum_g \sum_{g' \neq g} \sum_{l = \text{pipeline}} \sum_{i \in IL(l)} upoc_t Q_{i,g,g',l,t} \qquad \forall t \tag{29}$$

Finally, Eq. (30) calculates the ship operating costs based on the unit operating costs for maritime transportation ($usoc_t$), the time required to deliver the hydrogen and the cargo:

$$SOC_t = \sum_g \sum_{g' \neq g} \sum_{l = \text{ship}} \sum_{i \in IL(l)} usoc_t \left(distance_{g,g'} Q_{i,g,g',l,t} \right) \qquad \forall t \tag{30}$$

3.3.5.6 Environmental impact

The calculation of the environmental impact of the SC requires the quantification of the life cycle inventory of emissions and feedstock requirements, that is, of all relevant inputs and outputs of materials and energy associated with the network operation. The life cycle inventory (LCI) can be expressed as a function of some continuous decision variables of the model. Specifically, it can be calculated from the production rates at the plants ($PR_{i,g,p,t}$), and hydrogen flows ($Q_{i,g,g',l,t}$) as follows:

$$\begin{aligned} LCI_b = & \sum_i \sum_g \sum_p \sum_t PR_{i,g,p,t} \ (\omega_b^{Pr} \ + \ \omega_b^{St}) \\ & + \ \sum_i \sum_g \sum_{g' \neq g} \sum_{i \in IL(l)} \sum_t Q_{i,g,g',l,t} \ \omega_b^{Tr} \qquad \forall b \end{aligned} \tag{31}$$

The first term of Eq. (31) represents the emissions associated with the manufacture and storage of hydrogen. This term accounts for the emissions released during the extraction, processing and delivery of raw materials to the production facilities, the production of hydrogen itself, and its compression or liquefaction. The second term accounts for the emissions associated with the transportation of hydrogen between sub-regions. ω_b^{Pr}, ω_b^{St}, and ω_b^{Tr} denote the life cycle inventory entries (*i.e.*, emissions released to the environment or resource taken from the ecosphere) associated with chemical b per reference flow of activity. In the production and storage of hydrogen, the reference flow is one unit of main product produced/stored. In the transportation tasks, the reference flow is one unit of mass transported one unit of distance.

The damages in impact category d (DAM_d) are calculated from the life cycle inventory and the corresponding damage factors ($\upsilon_{b,d}$) as follows:

$$DAM_d = \sum_b \upsilon_{b,d} \cdot LCI_b \tag{32}$$

In this work, we assess the environmental performance of the hydrogen SC by means of the following 8 environmental LCA indicators: damage to human health caused by carcinogenic substances (CS), damage to human health caused by respiratory effects (RE), damage to human health caused by climate change (CC), damage to human health caused by ozone layer depletion (OLD), damage to ecosystem quality caused by ecotoxic substances (ES), damage to ecosystem quality caused by acidification and eutrophication (AE), damage to minerals (DM), and damage to fossil fuels (DFF). All these impacts are determined according to the Eco-indicator 99 methodology (ISO, 1997–2000).

The overall formulation can be finally expressed in compact form as follows:

$$
\begin{aligned}
(M) \qquad & \min_{y,Y,N} \{TDC(y,Y,N); DAM_1(y,Y,N), \ldots, DAM_D(y,Y,N)\} \\
s.t. \qquad & Eqs.\ (1) - \ (32) \\
& y \in \mathbb{R}, \quad Y \in \{0,1\}, \quad N \in \mathbb{Z}^+
\end{aligned}
\tag{33}
$$

Here, y denotes the continuous variables of the problem (capacity expansions, production rates, inventory levels and materials flows), Y represents the binary variables (*i.e.*, establishment of transportation links), and N are the integer variables denoting the number of plants, storage facilities and transport units.

Model (M) can be solved by any multi-objective optimization method (see Marler and Arora, 2004). In this work we obtain Pareto solutions using the ε-constraint method (Ehrgott, 2000a).

3.4 Proposed approach: combined use of multi-objective optimization and PCA

The multi-objective model presented above accounts for the simultaneous minimization of several criteria. The need to consider multiple environmental objectives represents a major obstacle, as it increases the problem complexity in terms of calculation and analysis of the Pareto set. To ameliorate these difficulties, we propose to use PCA (Jackson, 2003; Krzanowski, 2000) with the aim to identify and uncover relationships between objectives. Particularly, the outcome of the PCA will be utilized to eliminate redundant environmental impacts, thereby facilitating the visualization and interpretation of the solution space.

PCA is a multivariate technique that allows to identify inter-related variables and transform them into a smaller set of uncorrelated variables, known as principal components (PCs), which consist of a convex combination of the original variables. PCs are ordered according to the amount of variance they explain. Henceforth, the first j PCs can be used to select a subset of the original variables while still retaining most of the variance existing in the full-space problem.

Different methods based on PCA have been proposed so far for identifying a subset of uncorrelated variables from a wider set of correlated variables (see Zuur et al. (2007) for more details). We propose herein to take advantage of this particular application of PCA for identifying redundant environmental objectives in the multi-objective MILP formulation, with the final aim to reduce its dimensionality and enhance our understanding about its solution space.

Gutiérrez et al. (2010) explored the use of PCA to study relationships among different

LCA metrics in the context of waste water treatment plants. 7 different indicators (eutrophication, EU; acidification potential, AP; abiotic depletion; ADP, global warming, GW; photochemical oxidant formation potential, POFP; ozone layer depletion, OD; and terrestrial mycotoxicity potential, TETP) were considered in the study. After applying PCA, strong correlations were observed between ADP, GW and POFP, and between AP, OD and TETP. They also applied PCA and LCA to study mussel cultivation and waste electrical and electronic equipment.

To the best of our knowledge, Deb and Saxena (2005) were the first to use PCA in multi-objective optimization. They proposed a procedure to identify redundant objectives from the outcome of a PCA study performed on a set of feasible points of a multi-objective problem. Particularly, they developed some heuristic rules to reduce the dimensionality of the problem based on the components of the eigenvectors of the correlation matrix. The authors claimed that their method is able to provide good results for large-dimensional problems with up to 30 objectives. In this work, we apply a similar strategy to the multi-objective design of hydrogen infrastructures.

Our approach is as follows. First we generate a set of solutions of the original MILP (*i.e.*, the model with the whole set of objectives O) by using any available multi-objective algorithm. Assume that, the ε-constraint method is employed to generate a set of Pareto solutions R, that is, an $|R|$-by-$|O|$ matrix. Prior to the application of PCA we normalize this matrix by dividing each objective function value per the maximum one as follows:

$$nF_{o,r} = \frac{F_{o,r}}{\overline{F_o}} \tag{34}$$

where $F_{o,r}$ denotes the original value of the objective function to be normalized, and $\overline{F_o}$ is the maximum value of objective o over all the solutions (*i.e.*, $\overline{F_o} = \max\limits_{r=1,\ldots,|R|} \{F_{o,r}\}$). The data set is next standardized so as to make its centroid equal to zero. This is done by subtracting the mean of each column from each data point in the matrix and dividing the result by the standard deviation of the corresponding column. We next calculate the eigenvalues λ and eigenvectors X of the standardized correlation matrix. We then apply Kaiser-Guttman rule (Guttman, 1954; Kaiser, 1960) and exclude from the analysis the eigenvalues that are less or equal to 1. After sorting the remaining eigenvalues in a descendant order, the cumulative explained variance of the first j principal components (G_j) is determined as follows:

$$G_j = \sum_{e=1}^{j} \frac{\lambda_e}{\sum\limits_{e=1}^{|\lambda|} \lambda_e} \tag{35}$$

Deb and Saxena (2005) suggested to define a threshold cut (CUT), typically 95%, and keep for the PCA the eigenvalues with cumulative explained variance below it. Hence, the principal components (*i.e.*, the first j eigenvalues with $G_j \leq$ CUT and the corresponding eigenvectors) of the data set are then selected for the analysis. Note that the appropriate value of CUT is case-dependent and may be greater or lower than 95%. We propose to use the graphical Cattell Scree test (Cattell, 2009) for determining the number of components to retain in PCA. The Scree test involves plotting the eigenvalues in a descendant order against their sequence numbers and determining the changeover from a steep slope to a levelling off. The components to the right of this break should be dropped from further analysis. The lack of any clear break or multiple break points would make impossible to use the Scree test. In this case another approaches to determining the number of components should be used (see Jackson, 1993).

The next step is analyzing the eigenvectors (typically referred to as loadings). The elements of the loadings are employed to identify conflicts among objectives. Particularly, the first elements of the loadings denote the contribution of the first objective in the principal components, the second ones denote the contribution of the second function, and so on. The most-positive value x^+ corresponds to the objective that causes the maximum increase in the principal component, whereas the most-negative element x^- denotes the function causing the largest decrease. Hence, the objectives corresponding to the most-positive and most-negative elements of the factor loading are regarded as the most conflicting objectives within that principal component.

Different methods are available in the literature for reducing the dimensionality of a data set using PCA. Without loss of generality, we follow herein the heuristic procedure suggested by Deb and Saxena (2005). In the context of our problem, this method enables us to identify conflicts and redundancies among LCA metrics. This information will be employed to remove redundant objectives from the analysis. Technical details about this strategy can be found elsewhere, and are summarized in Figure 3.15 at the end of the chapter. Note that other procedures based on the analysis of the correlation matrix could be used for reducing the dimensionality of the problem (see Goel et al., 2007).

We should clarify at this point that in this work the emphasis is placed on the use of PCA in the post-optimal analysis of the MILP. Note, however, that it is possible to use the same strategy in an iterative manner. That is, applying PCA for identifying redundant objectives and then resolving the model in the reduced space of objectives and repeating the overall procedure until a stopping criterion is satisfied.

3.5 Case study

3.5.1 Dataset and assumptions

The capabilities of the proposed methodology are illustrated through its application to a case study based on Spain. A planning horizon of 8 years is defined. We consider 19 potential locations for the establishment of production and storage technologies that are defined according to the administrative divisions (autonomous communities) of Spain.

The corresponding hydrogen demand, which has been determined assuming a market penetration of 12.9% in 2040 and 36% in 2050 in the current energy system based on fossil fuels, is displayed in Table 3.1 (HyWays, 2004; Seydel and Wietschel, 2007; HyWays, 2010). A minimum demand satisfaction level of 90% is set. Distances between sub-regions were calculated considering the capitals of the autonomous communities and the main roads or maritime routes connecting them. These data are listed in Table 3.2.

Table 3.1: Hydrogen demand, kg/day

Sub-region	Autonomous community	2040–2042	2042–2044	2044–2046	2046–2048
G01	Andalusia	723,144	942,895	1,162,646	1,382,397
G02	Aragon	125,896	164,154	202,412	240,669
G03	Principality of Asturias	83,833	109,308	134,784	160,259
G04	Balearic Islands	114,722	149,585	184,447	219,309
G05	Canary Islands	240,767	313,932	387,097	460,262
G06	Cantabria	50,214	65,473	80,732	95,991
G07	Castile and León	233,435	304,372	375,309	446,245
G08	Castile-La Mancha	209,819	273,580	337,340	401,101
G09	Catalonia	680,427	887,197	1,093,966	1,300,736
G10	Ceuta	8,772	11,438	14,104	16,769
G11	Valencian Community	458,320	597,595	736,871	876,146
G12	Extremadura	103,761	135,293	166,824	198,355
G13	Galicia	236,528	308,405	380,282	452,159
G14	La Rioja	29,883	38,963	48,044	57,125
G15	Madrid	566,728	738,947	911,166	1,083,384
G16	Melilla	5,263	6,863	8,462	10,062
G17	Region of Murcia	137,293	179,014	220,735	262,456
G18	Foral Community of Navarre	63,563	82,878	102,194	121,510
G19	Basque Country	170,566	222,397	274,229	326,061

Table 3.2: Distances between sub-regions, km

	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19
G01	0	854	852	943	1,550	842	586	493	998	223	658	188	958	919	528	440	534	981	884
G02	854	0	608	574	2,403	401	438	414	307	1,021	312	668	838	176	324	968	579	180	258
G03	852	608	0	1,177	2,371	194	294	522	910	1,044	804	640	335	440	447	1,177	851	502	406
G04	943	574	1,177	0	2,448	970	1,007	724	267	984	355	1,040	1,407	745	699	811	422	748	827
G05	1,550	2,403	2,371	2,448	0	2,389	2,133	2,024	2,541	1,586	2,201	1,735	2,505	2,466	2,075	1,981	2,036	2,528	2,431
G06	842	401	194	970	2,389	0	249	534	704	1,062	709	658	500	234	452	1,189	862	254	162
G07	586	438	294	1,007	2,133	249	0	287	739	806	569	402	448	269	212	943	616	331	235
G08	493	414	522	724	2,024	534	287	0	715	650	383	304	677	477	90	680	393	539	442
G09	998	307	910	267	2,541	704	739	715	0	1,181	352	968	1,140	477	623	1,008	619	481	559
G10	223	1,021	1,044	984	1,586	1,062	806	650	1,181	0	821	408	1,178	1,090	706	395	572	1,142	1,056
G11	658	312	804	355	2,201	709	569	383	352	821	0	698	967	491	357	648	259	494	573
G12	188	668	640	1,040	1,735	658	402	304	968	408	698	0	785	731	340	623	725	793	696
G13	958	838	335	1,407	2,505	500	448	677	1,140	1,178	967	785	0	669	600	1,398	1,004	731	634
G14	919	176	440	745	2,466	234	269	477	477	1,090	491	731	669	0	385	1,141	752	86	90
G15	528	324	447	699	2,075	452	212	90	623	706	357	340	600	385	0	737	404	448	351
G16	440	968	1,177	811	1,981	1,189	943	680	1,008	395	648	623	1,398	1,141	737	0	399	1,145	1,085
G17	534	579	851	422	2,036	862	616	393	619	572	259	725	1,004	752	404	399	0	755	760
G18	981	180	502	748	2,528	254	331	539	481	1,142	494	793	731	86	448	1,145	755	0	99
G19	884	258	406	827	2,431	162	235	442	559	1,056	573	696	634	90	351	1,085	760	99	0

The minimum and maximum production capacities of each technology are given in Table 3.3 (Almansoori and Shah, 2006; Kim et al., 2008; Padró and Putsche, 1999). Table 3.4 shows the capital cost of the production facilities considered in this study, whereas the associated production costs are presented in Table 3.5 (Almansoori and Shah, 2006; Kim et al., 2008; Padró and Putsche, 1999). The minimum and maximum storage capacities for liquid and compressed H_2 are 10,000 and 23,240,000 kilograms, respectively (Amos, 1998).

Table 3.3: Minimum and maximum production capacities of each technology, kg /day

	Technologies		
	SMR	CG	WE
Minimum production capacity	10,000	10,000	10,000
Maximum production capacity	480,000	480,000	50,000

Table 3.4: Capital cost of production facilities, million €

	Hydrogen form	
	Liquefied	Compressed
SMR	465	329
CG	832	670
WE	97	47

Table 3.5: Hydrogen production cost, €/kg

	Hydrogen form	
	Liquefied	compressed
SMR	1.33	0.82
CG	1.49	0.92
WE	2.93	2.28

The unit storage cost is €0.004 /(kg·day) for liquefied hydrogen, and €0.066 /(kg·day) for compressed hydrogen (Amos, 1998). The capital and operating cost parameters of trucks and railroad cars can be found in Table 3.6 (Almansoori and Shah, 2006; Amos, 1998). The capital cost associated with the establishment of pipelines is €708,673 per km (Almansoori and Shah, 2006; Amos, 1998). The daily transportation cost of hydrogen using pipelines is €0.0576732 per kg of H_2 (Amos, 1998). The model parameters taken from the studies corresponding to United Kingdom or US were converted from local currencies to euros using the average exchange rates (IMF, 2009). Moreover the consumer price index (INE, 2009) and a correcting factor of 0.8 were applied to the parameters corresponding to United Kingdom. The hourly freight cost associated with maritime transport is €0.00115 per kg of H_2 (Transmar S. L., 2009).

The minimum flow rate of each transportation mode is assumed to be equal to the minimum capacity of the corresponding transportation mode, whereas the maximum flow rate for any type of transportation mode is 96,000,000 kg per day (Almansoori and Shah, 2006).

The LCA impact was quantified according to the Eco-indicator 99 methodology, using the hierarchist perspective and average weighting (H/A). The emissions and feedstock requirements associated with SMR, CG and WE technologies were calculated according to the studies by Spath and Mann (2001), Koroneos et al. (2005) and Spath and Mann (2004), respectively. The work of Koroneos et al. (2005) considers only the emissions associated with coal gasification, and neglects those taking place during coal extraction and transportation from mines to process plants. We filled these data gaps using the Ecoinvent database (Ecoinvent Centre, 2010). For the transportation of coal, we made the assumption that coal is transported by diesel train, and assumed an average distance between the coal mine and the CG facility of 50 km. The impact of the transportation tasks using trucks, rail cars, onshore pipelines and ships was also retrieved from Ecoinvent. The LCA impacts due to the production, storage and transportation technologies are shown in Table 3.7.

Table 3.6: Parameters used to calculate the capital and operating cost for different terrestrial transportation modes

	Tanker truck	Tube trailer	Railway tank car	Railway tube car
Average speed (km /h)	55	55	45	45
Capacity (kg /trip)	4,082	181	9,072	454
Cost of establishing transportation mode (€)	434,236	217,118	434,236	260,541
Availability of transportation mode (h /day)	18	18	12	12
Driver wage (€/h)	19.97	19.97	19.97	19.97
Fuel economy (km /L)	3.58	3.58	10.13	10.13
Fuel price (€/L)	1.01	1.01	0.24	0.24
General expenses (€/day)	7.14	7.14	5.95	5.95
Load/unload time of product (h /trip)	2	2	12	12
Maintenance expenses (€/km)	0.085	0.085	0.054	0.054

Table 3.7: Values of the environmental indicators, ecopoints

	Unit	CS	RE	CC	OLD	ES	AE	DM	DFF
Production									
SMR	kg of H_2	9.0860×10^{-5}	7.8775×10^{-2}	6.4771×10^{-2}	0	0	6.2449×10^{-3}	7.1070×10^{-6}	3.9662×10^{-1}
CG	kg of H_2	8.4794×10^{-4}	6.7636×10^{-1}	2.4414×10^{-1}	4.6830×10^{-7}	3.1598×10^{-4}	3.1549×10^{-2}	1.2621×10^{-4}	7.1612×10^{-2}
WE	kg of H_2	0	5.4196×10^{-1}	5.3012×10^{-3}	0	0	2.5868×10^{-3}	1.4642×10^{-4}	9.7977×10^{-3}
Storage									
Liquefied H_2	kg of H_2	5.6015×10^{-2}	1.4761×10^{-1}	3.2377×10^{-2}	8.7914×10^{-6}	1.6468×10^{-2}	1.2151×10^{-2}	4.6896×10^{-3}	1.2410×10^{-1}
Compressed H_2	kg of H_2	1.2448×10^{-2}	3.2802×10^{-2}	7.1949×10^{-3}	1.9536×10^{-6}	3.6595×10^{-3}	2.7003×10^{-3}	1.0421×10^{-3}	2.7577×10^{-2}
Transportation									
Truck	tkm	6.9355×10^{-7}	7.3664×10^{-6}	1.3999×10^{-6}	1.0963×10^{-9}	6.5762×10^{-7}	9.2240×10^{-7}	1.1688×10^{-7}	1.2057×10^{-5}
Railway car	tkm	4.3810×10^{-8}	5.0619×10^{-7}	1.4969×10^{-7}	3.8708×10^{-11}	1.1919×10^{-8}	4.2717×10^{-8}	2.2452×10^{-9}	5.1874×10^{-7}
Pipeline	tkm	1.8678×10^{-8}	4.8789×10^{-7}	3.0030×10^{-7}	3.2098×10^{-9}	1.0106×10^{-8}	7.9442×10^{-8}	6.2837×10^{-9}	2.8478×10^{-6}
Ship	tkm	2.0453×10^{-8}	6.4776×10^{-7}	5.8493×10^{-8}	3.2964×10^{-11}	3.6928×10^{-8}	7.6013×10^{-8}	1.2184×10^{-9}	4.4695×10^{-7}

3.5.2 Results and discussion – MOO Problem

To simplify the calculations, we generated the Pareto solutions by solving a set of bi-criteria models in which the cost was optimized against each single LCA impact separately. Note that each of the points calculated in this way is guaranteed to be weakly Pareto optimal in the original search space. Each of these bi-criteria problems was written in GAMS (Rosenthal, 2008) and solved with the MILP solver CPLEX 12.0 on a HP Compaq DC5850 desktop PC with an AMD Phenom 8600B, 2.29 GHz triple-core processor, and 2.75 Gb of RAM.

We ran the ε-constraint method for each of these bi-criteria models, obtaining finally 41 Pareto solutions. The CPU time required to calculate a single Pareto solution varies from 2 minutes to 4 hours. Figures 3.2 to 3.9 show the corresponding bi-criteria Pareto sets. Particularly, in each plot we show for each Pareto point, the normalized values of all the environmental metrics. The normalization has been performed by dividing the environmental performance by the maximum impact value over all Pareto solutions. The points highlighted in bold correspond to the values of the impact being optimized (*i.e.*, the one being minimized in the bi-criteria model), whereas the remaining ones represent the environmental performance of each solution in the remaining categories. As observed in Figures 3.3, 3.5, 3.6 and 3.8, the environmental metrics tend to behave in a similar manner, that is, when one of them is minimized, the remaining ones are also reduced.

In contrast, in Figures 3.2, 3.4, 3.7 and 3.9, as we move from right to left we observe that certain metrics are monotonically increasing while some others go up and down depending on the cost value.

Particularly, three regions can be distinguished in Figure 3.2 that differ in the production and storage technologies as well as in the schemes of hydrogen delivery. In interval AB, hydrogen is produced *via* SMR , and the solutions mainly differ in the amounts of hydrogen compressed and liquefied. From A to B, the amount of compressed hydrogen increases, since this storage technology causes less impact in the considered environmental metrics. WE facilities appear after point B, and its number gradually increases as we approach point D. Note that water electrolysis coupled with wind turbines produces much more solid particles than SMR facilities (28.7 g/kg of H_2 in the case of WE and 2 g/kg of H_2 in the case of SMR). Because of this, the RE values increase gradually after point B. Moreover, WE plants require more concrete and steel than SMR facilities. Therefore, the inclusion of WE plants in the hydrogen network increases minerals consumption and consequently the DM impact. This effect is only observed in the interval CD.

In Figures 3.3, 3.5, 3.6 and 3.8 all the configurations operate exclusively with SMR facilities. These solutions differ only in the amounts of transported hydrogen and its physical form. As we move from A to B, the amount of compressed H_2 increases. Similarly, we obtain more decentralized SC requiring less transportation tasks. In point

B (*i.e.*, minimal impact solution), the compressed hydrogen demand is fully covered with the internal production facilities located within each region. As a result, all curves monotonically decrease from A to B, since the associated structural changes reduce simultaneously the impact in all damage categories.

In Figures 3.4 and 3.7, three regions with different SC topologies are identified. In the interval AB, liquefied hydrogen is produced *via* technologies SMR and WE. The steep increase in impacts RE and DM is due to the increasing share of WE. The SC configurations placed on the right hand side of point B produce hydrogen in both physical forms: compressed and liquefied. These designs open only WE facilities. The decrease in the values of RE and DM is due to the increase in the percentage of compressed H_2, which is the storage technology with the best environmental performance in all of the environmental metrics.

Coal gasification appears only in the interval AB on Figure 3.9. It causes the raise of all metrics except DFF (the one being optimized). After point B, only WE is selected, and solutions in the interval CD differ in both the amount of compressed and liquefied hydrogen and scheme of hydrogen delivery. This shift from the combined use of CG and SMR to WE technology decreases all of the environmental metrics in point B except DFF.

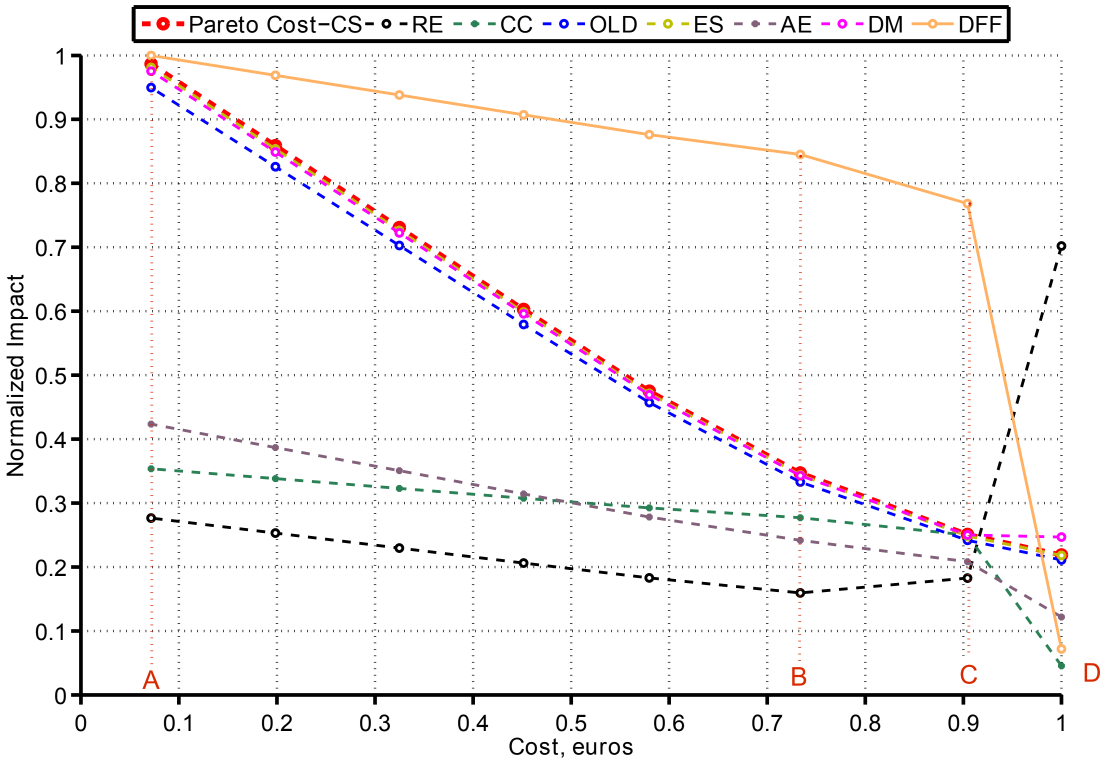


Figure 3.2: Pareto set of solutions between TDC and CS (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

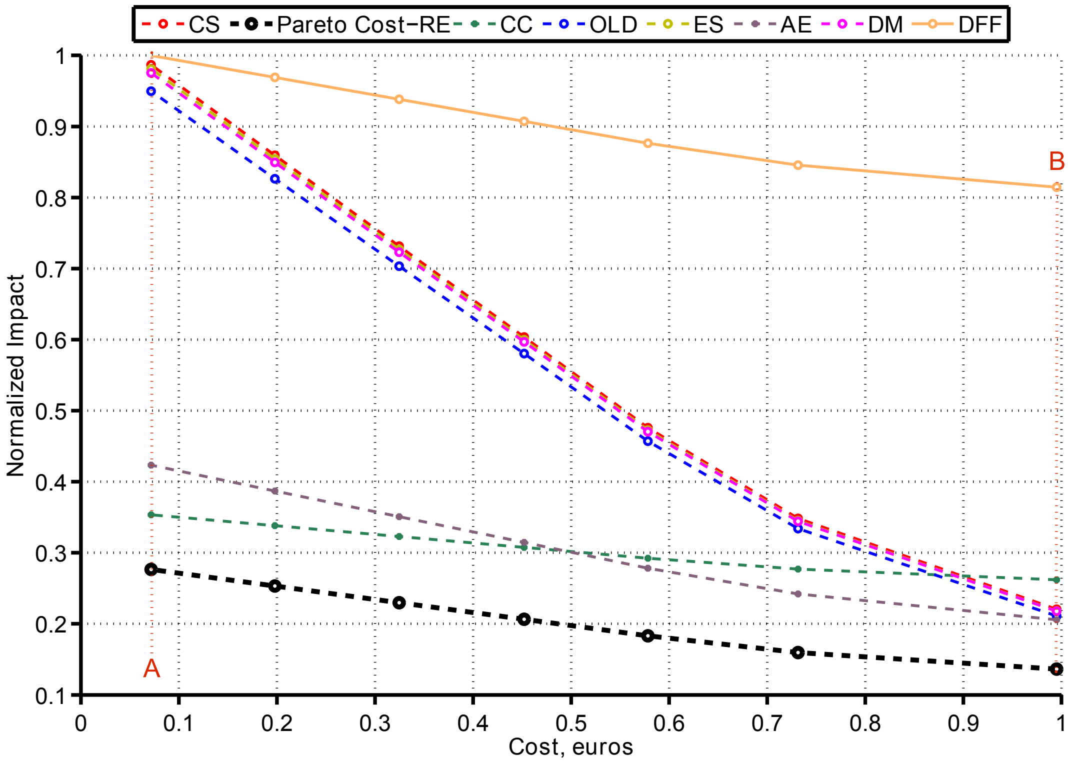


Figure 3.3: Pareto set of solutions between TDC and RE (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

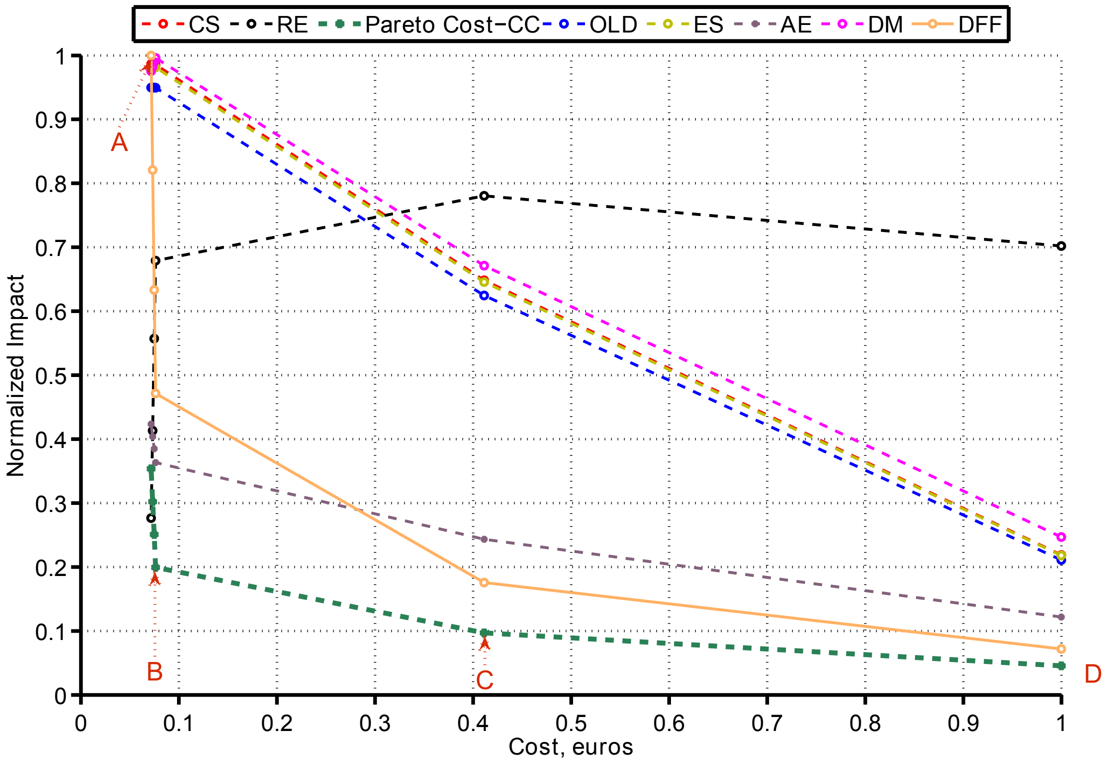


Figure 3.4: Pareto set of solutions between TDC and CC (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

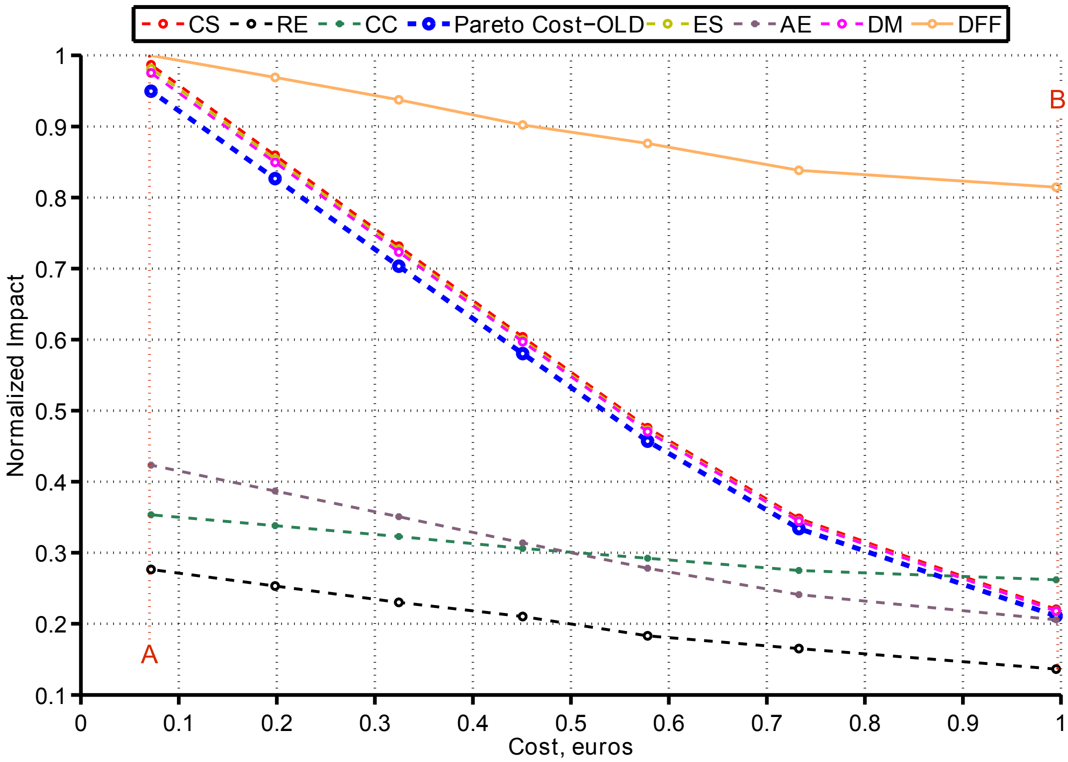


Figure 3.5: Pareto set of solutions between TDC and OLD (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

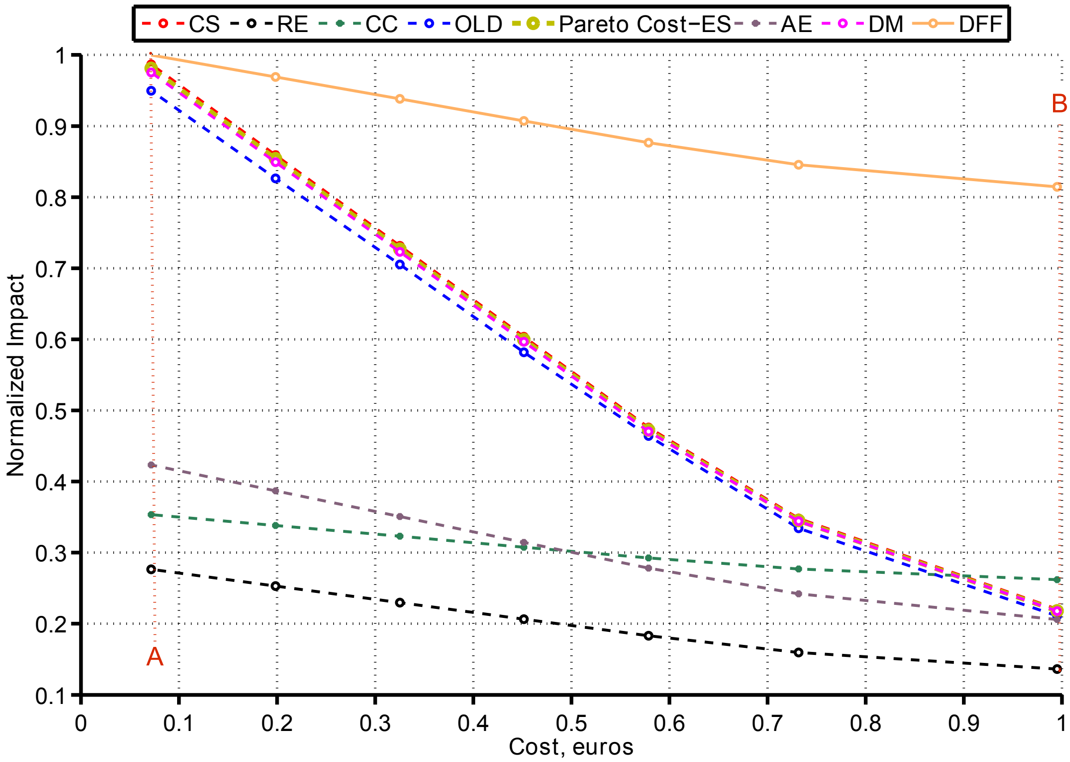


Figure 3.6: Pareto set of solutions between TDC and ES (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

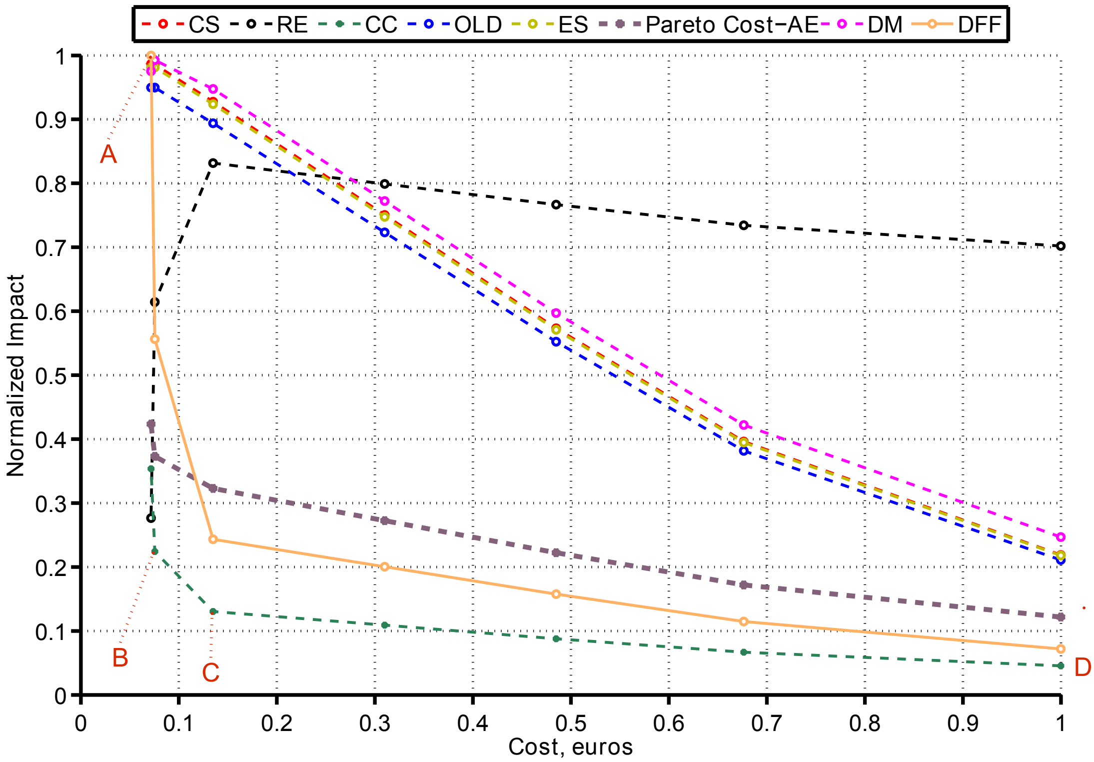


Figure 3.7: Pareto set of solutions between TDC and AE (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

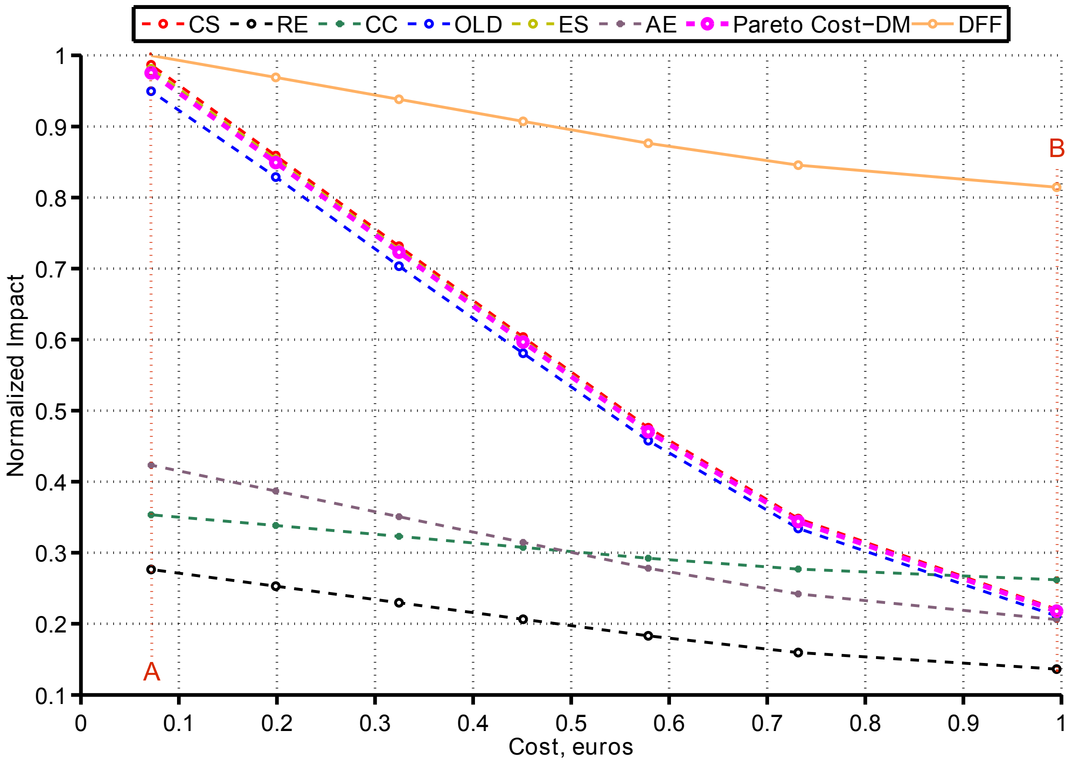


Figure 3.8: Pareto set of solutions between TDC and DM (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

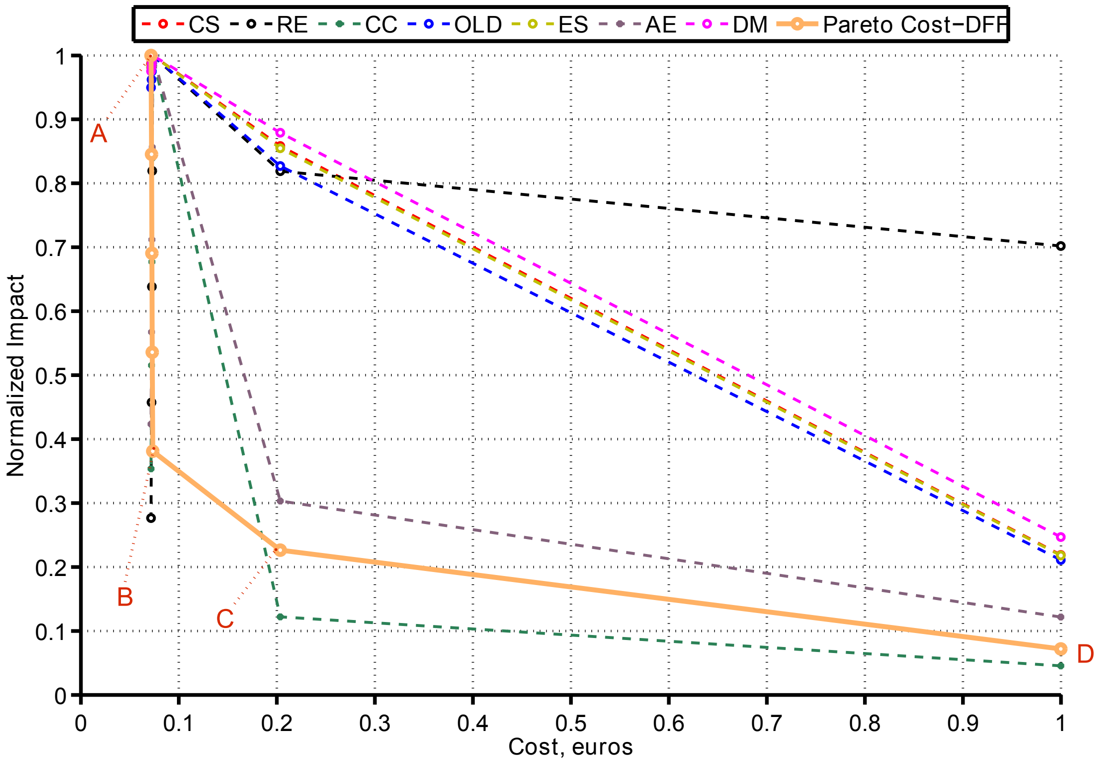


Figure 3.9: Pareto set of solutions between TDC and DFF (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

Figures 3.10 to 3.12 summarize the main strategic decisions associated with the extreme solutions of the Pareto sets in Figures 3.2 to 3.9 . Interestingly, the minimization of some environmental metrics leads to the same SC configuration. More precisely, two minimum impact SC designs are identified: one for the minimization of CS, CC, AE, and DFF and another one for the minimum RE, ES, OLD, and DM solutions.

In the minimum cost solution, all production facilities are based on SMR and produce liquefied hydrogen that is stored in cryogenic spherical tanks. The SC structure is quite decentralized, and many regions (ten in total) fulfill their demand by domestic SMR plants, except few of them that import liquid H_2 from neighboring communities. Short-distance terrestrial transportation is carried out with trucks, whereas for middle-distance the model selects railroad. Melilla and the Balearic and Canary Islands import hydrogen *via* freighted ships. The demand in Ceuta is fully satisfied by the domestic SMR production facilities.

Figure 3.11 presents the minimum CS, CC, AE, and DFF network. In this SC configuration, compressed hydrogen is produced *via* WE. The SC is pure decentralized, and does not require any transportation. The SC configuration with minimum RE, ES, OLD, and DM also produces compressed hydrogen but using SMR technology instead of WE. Figure 3.12 shows the corresponding SC structure. As observed in this network, the hydrogen flows between sub-regions are all zero.

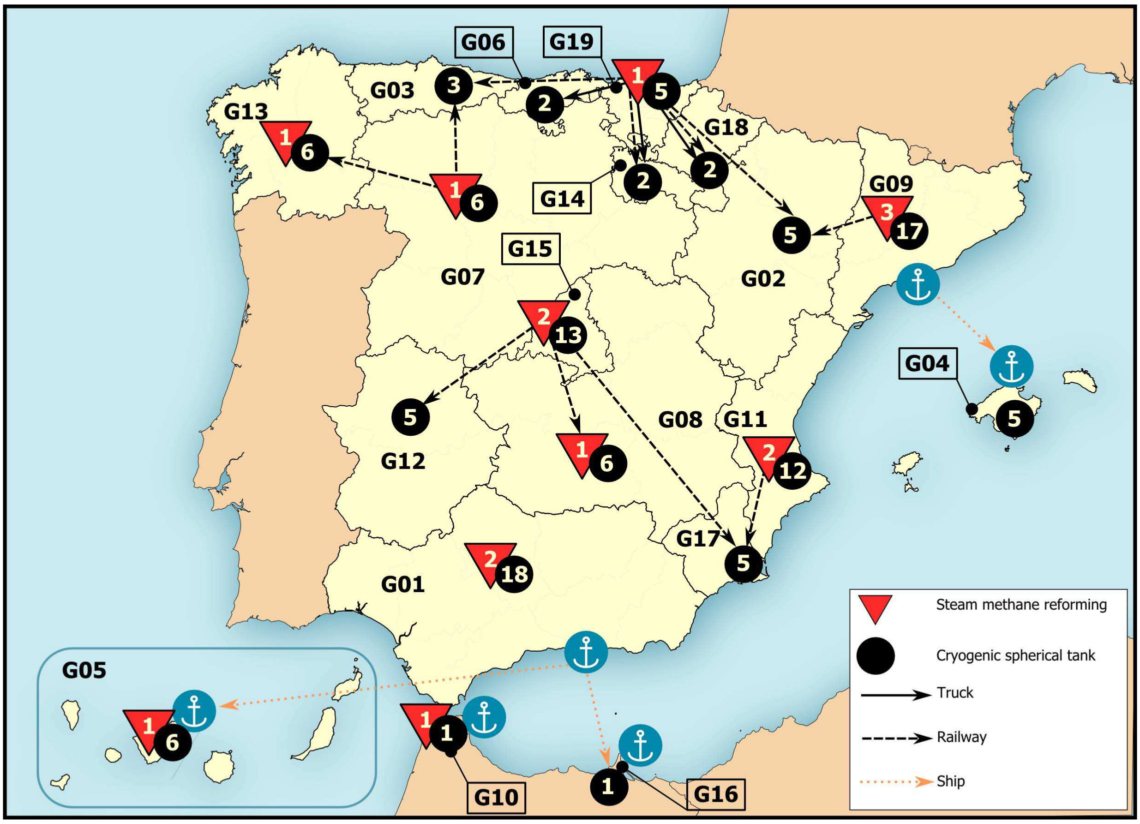


Figure 3.10: SC configuration for the solution with minimum TDC

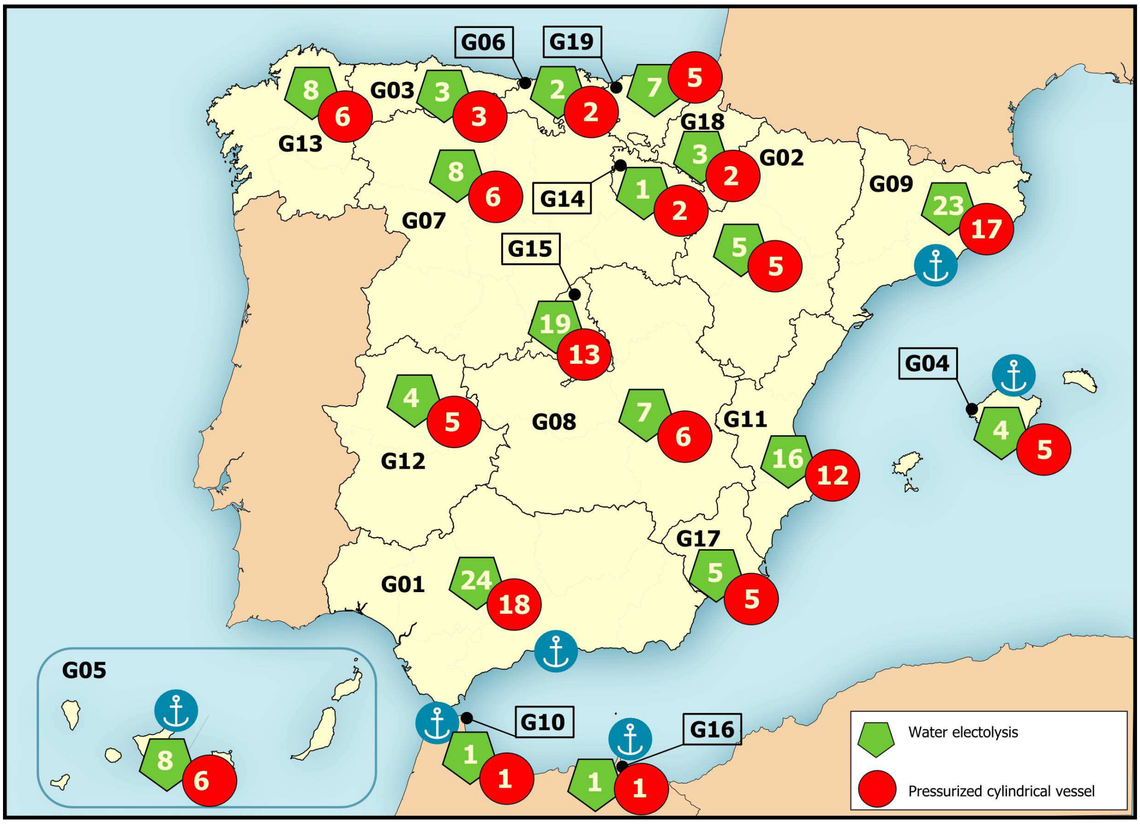


Figure 3.11: SC configuration for the solutions with minimum CS, CC, AE and DFF

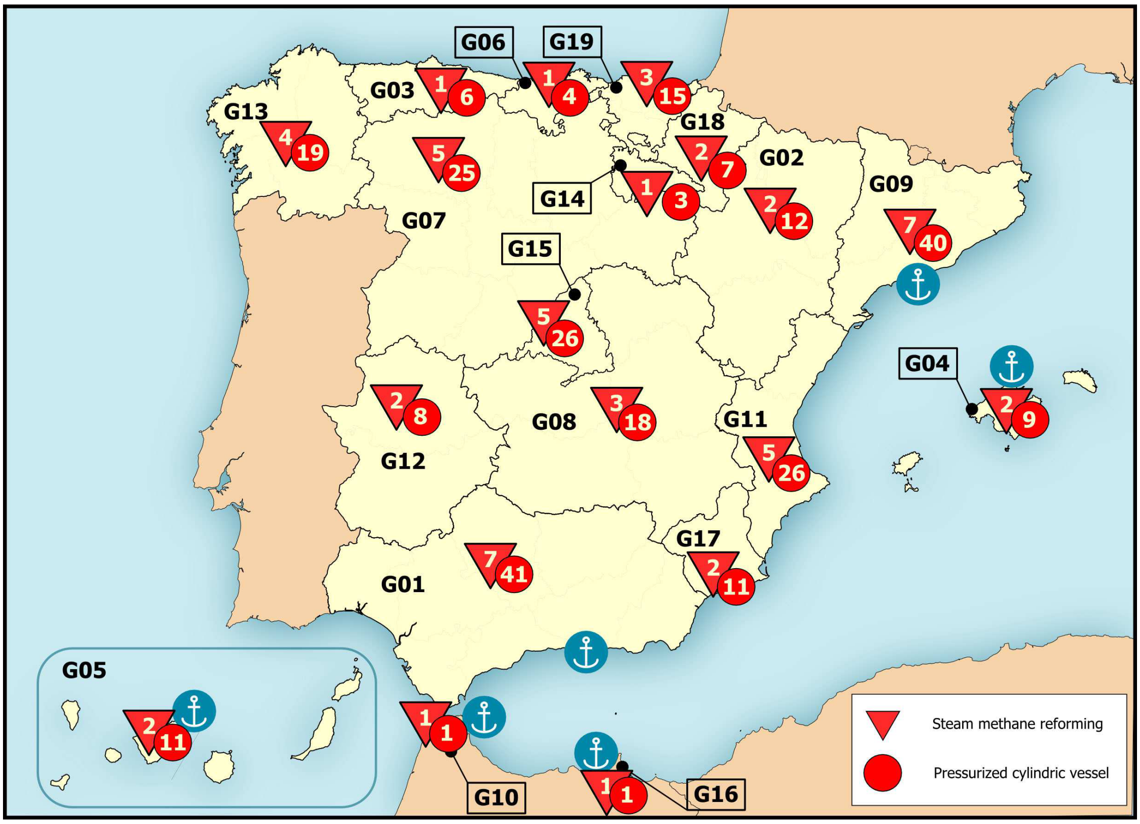


Figure 3.12: SC configuration for the solution with minimum RE, ES, OLD and DM

We next applied PCA to these solutions. The correlation matrix is shown in Table 3.8. The closer the elements of the matrix are to -1 or 1, the stronger is the correlation between the corresponding objectives. It can be seen that metrics CS, ES, OLD and DM are somehow equivalent, whereas AE is highly correlated with CC. The eigenvectors and eigenvalues of the correlation matrix are presented in Table 3.9, where the principal components are arranged in a descent order. The most-negative and most-positive elements of the principal components are highlighted in bold font. The conflicting objectives are written below the cumulative explained variance of every principal component.

Table 3.8: Correlation matrix

	CS	RE	CC	OLD	ES	AE	DM	DFF
CS	1	0.4400	0.4007	0.9995	1.0000	0.6729	0.9988	-0.0375
RE		1	0.0735	0.4517	0.4435	0.3886	0.4792	-0.8966
CC			1	0.4235	0.4027	0.9237	0.377	0.2743
OLD				1	0.9996	0.6938	0.9984	-0.0453
ES					1	0.6753	0.9989	-0.0408
AE						1	0.6621	0.0165
DM							1	-0.0847
DFF								1

Table 3.9: PCA results

	Principal components							
	1	2	3	4	5	6	7	8
CS	-0.4290	0.0188	0.2517	0.3581	-0.3265	0.5347	-0.0827	-0.4739
RE	-0.2436	-0.5752	-0.2668	0.3797	0.4647	-0.1424	-0.3915	-0.0696
CC	-0.2529	0.3791	-0.6238	0.1948	-0.2479	-0.4454	0.1347	-0.2955
OLD	-0.4322	0.0189	0.2230	-0.6102	-0.1216	-0.3151	-0.5086	-0.1337
ES	-0.4295	0.0171	0.2478	0.3595	-0.2663	-0.2309	0.0645	0.7044
AE	-0.3659	0.1852	-0.4966	-0.2990	0.1950	0.5695	-0.0363	0.3634
DM	-0.4293	-0.0165	0.2498	-0.1412	0.4973	-0.1322	0.6567	-0.1925
DFF	0.0571	0.6999	0.2370	0.2805	0.4949	-0.0298	-0.3553	0.0012
Eigenvalues	5.0726	1.8888	1.0386	1.29×10^{-4}	4.76×10^{-5}	4.06×10^{-5}	5.07×10^{-6}	8.02×10^{-7}
Rank	1	2	3					
Explained variance	63.4081	23.6095	12.9821					
Cumulative explained variance	63.4081	87.0176	99.9997					
Conflicting objectives	OLD, DFF	RE, DFF	CC					
Redundant objectives	CS, ES, AE, DM							

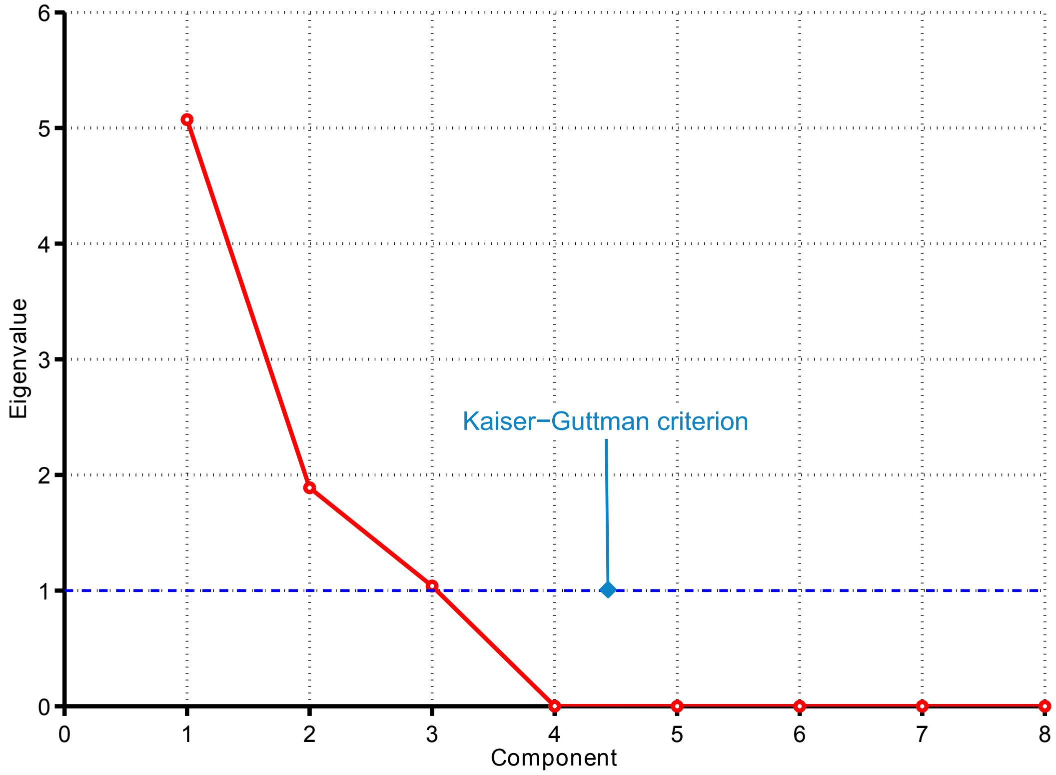


Figure 3.13: Scree plot

We performed a graphical Scree test to decide the number of principal component to be kept for further analysis. As observed in Figure 3.13, the levelling curve starts from the forth principal component. However, the forth and subsequent components do not satisfy the Kaiser-Guttman rule, so we analyzed only the eigenvectors corresponding to the first three principal components with a cumulative explained variance of almost 100%. As observed, four environmental indicators can be excluded from the set of objectives following the heuristic rule proposed by Deb and Saxena (2005).

Figure 3.14 shows the bi-dimensional plots representing the loads of the environmental objectives projected onto the sub-spaces of the first three principal components. The metrics selected as redundant are highlighted in blue color, whereas the conflicting objectives are shown in black.

As observed, the correlation matrix suggests that metrics CS, ES, DM and OLD are equivalent. Particularly, the correlation coefficients corresponding to these objectives are higher than 0.9. These metrics have very similar loads in the first three principal components. Because of this, their projections onto the axes shown in Figure 3.14 are quite close. One objective, namely OLD, can be retained while the remaining ones can be excluded from the analysis.

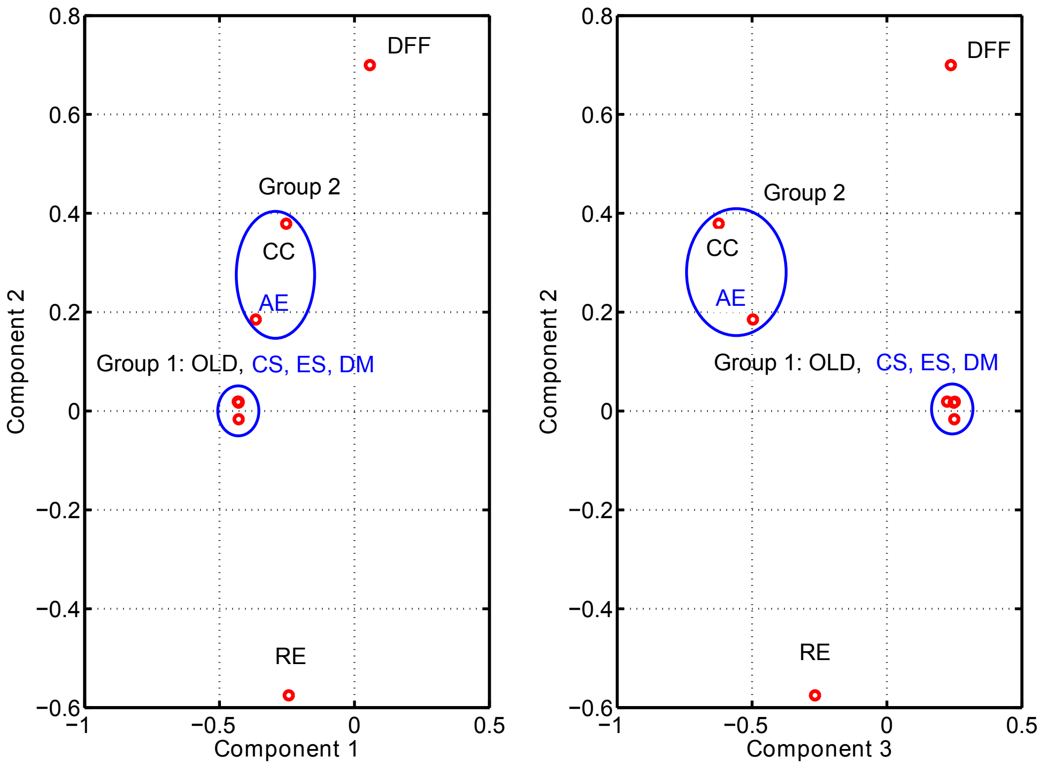


Figure 3.14: PCA results in the bi-dimensional spaces (CS — damage to human health caused by carcinogenic substances; RE — damage to human health caused by respiratory effects; CC — damage to human health caused by climate change; OLD — damage to human health caused by ozone layer depletion; ES — damage to ecosystem quality caused by ecotoxic substances; AE — damage to ecosystem quality caused by acidification and eutrophication; DM — damage to minerals; DFF — damage to fossil fuels)

The correlation between AE and CC is also high, and the proposed analysis suggests to omit impact AE. In Figure 3.14, these two environmental categories lie in the same quadrant and differ mainly in their projections on the second principal component. Furthermore, RE and DFF are conflicting according to the correlation matrix, since they show a very high negative correlation coefficient.

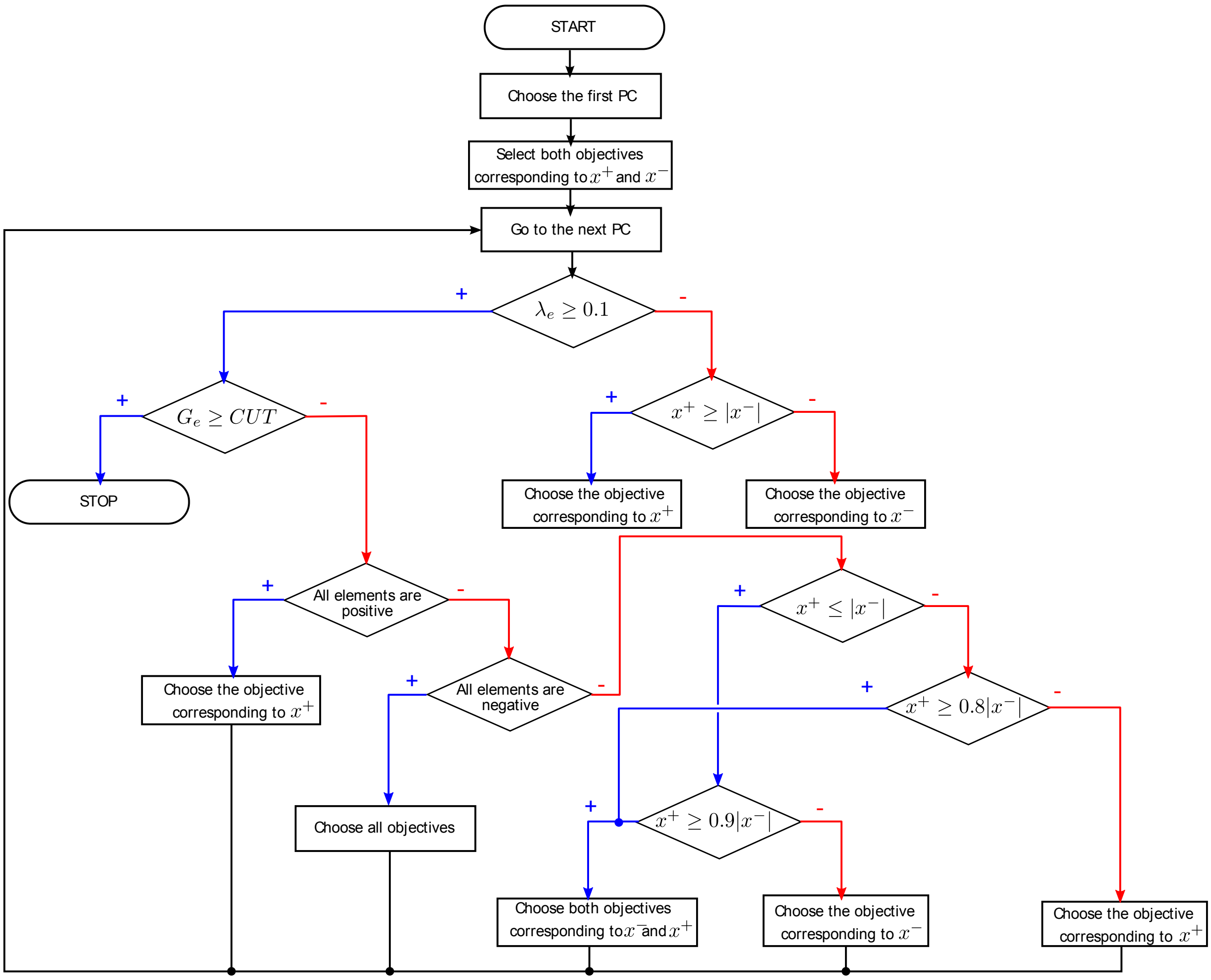


Figure 3.15: Sche me of the PCA procedure (Deb and Saxena, 2005), where x^+ and x^- are the most positive and the most negative elements of an eigenvector), respectively.

3.6 Conclusions

This work has proposed a systematic approach for the multi-objective design of hydrogen SCs for vehicle use that considers simultaneously several life cycle environmental criteria along with an economic performance metric. The environmental performance of the SC has been measured by independently evaluating eight life cycle assessment metrics widely used in the most common life cycle environmental impact methodology (*e.g.*, EcoIndicator 99). These in turn are used to calculate the damage caused by each specific SC according the different damage categories.

The proposed methodology comprises two main steps: in step one, a multi-objective MILP is constructed and a set of Pareto solutions are generated that represent the optimal trade-off between the all the pair combinations of economic and environmental objectives considered in the analysis. Since the analysis of the resulting designs proves to be a non-trivial task, a multi-variable statistical method (*i.e.*, PCA) is then applied to detect and omit redundant environmental indicators. This method can identify the environmental objectives that can be left out of the analysis without disturbing the main features of the solution space. The main novelties of the work lie on the use of eight standard life cycle environmental impact metrics and the aplication of PCA as dimensionality reduction technique to interpret the different optimal hydrogen supply chain designs arising from a multidimensional multi-objective environmental design problem. We have demonstrated the capabilities of this technique through a case study based on the future (potential) SC to be established in Spain.

In Section 3.1 the importance of environmental criteria in the evaluation of hydrogen supply networks is put into context. More specifically, an exhaustive literature review on the topic of hydrogen supply chain optimization revealed the narrow scope followed in the few works implementing multi-objective optimization to include environmental criteria. These works either did not consider life cycle environmental metrics or just implemented two objectives in their formulation, in which the optimal performance of a specific environmental metric might imply the deterioration of other environmental categories. Our approach attempts at overcoming these limitations by providing a full multi-objective framework embedding all possible environmental categories in the first instances, for then combining it with PCA into a single iterative framework that is able to uncover conflicting and redundant environmental criteria. Thus the resulting holistic approach provides an effective solution for the problem of visualizing highly complex multidimensional problems.

The mathematical formulation of this problem is introduced in Section 3.2, and then in Section 3.4 the post-optimality PCA analysis methodology is presented. Therefore, the combined use of MOO and PCA into a single integrated framework that is designed to be used in an iterative manner proves instrumental in uncovering redundant environmental metrics by analyzing their internal correlation structures. Finally a graphical test is presented as a useful tool to determine the appropriate threshold level to be applied for determining the number of principal components.

Note that this work does not include the production of hydrogen *via* biomass gasification since this specific framework was considered inappropriate to evaluate that technology for two main reasons:

- raw material and transportation environmental impacts would not be negligible and they would need to be modelled for a hypothetical case in Spain
- the uncertainties associated with the aforementioned biomass life cycle environmental impact calculations would not be properly covered in a deterministic framework

On the other hand, this approach presents a more realistic hydrogen demand profile accross different time periods than the one introduced in Chapter 2. Particularly, hydrogen is assumed to follow penetration rates according to similar trends projected in wider European research projects covering this topic. For the rest of the parameters, the same main assumptions and datasets as the ones introduced in Chapter 2 are used for comparison purposes.

In terms of the environmental performance measured according to the eight different life cycle categories, the results presented in Section 3.5 reveal two distinct groups of metrics:

- environmental metrics that simultaneously optimize the whole objective set
- environmental metrics whose optimization leads to dispar trends in the rest of the objectives

For all of them, the economic optimization solution presents a decentralized SMR based supply chain producing hydrogen in liquid form. Note that the major difference with the optimal economic performance solution presented in Chapter 2 is the decentralization of the supply chain design. This can be explained by the new more gradual and realistic demand profile introduced. Since the optimal decisions are taken in each period with the information available at that time, economies of scale are not fully exploited but the supply chain design is more flexible to introduce different technology mixtures. This analysis could easily form part of a more systematic structural uncertainty analysis for different demand profiles.

Interestingly, the optimization of the environmental metrics that simultaneously optimize the rest of the objectives (*i.e.*, RE, ES, OLD and DM) gives rise to a hydrogen SC configuration also decentralized but producing gaseous hydrogen stored in compressed form. This appears to respond for the negative environmental impacts of the energy and electricity used in the liquefaction step. The rest of the environmental objectives (*i.e.*, CS, CC, AE and DFF), produced a structure where decentralized water electrolysis (WE) plants using wind turbines were installed to produce gaseous hydrogen. As a result, the solutions obtained indicate that a decentralized structure is more advantageous in both SMR and WE hydrogen supply chains, while gaseous hydrogen is the most environmentally friendly alternative in any of the cases.

Note that WE is mainly selected for the reduction of only specific environmental metrics, but it might be the case that these metrics are considered to be the most important by decision-makers (*i.e*, climate change (CC)). The analysis reveals the methodology uncertainty associated to the use (or *bias*) of a single environmental metric. In fact, the solution minimizing cost and CC, although achieves a reduction in most other environmental metrics, it increases quite respiratory effects on humans (*i.e.* RE).

It is important to mention that limitations of these approach may lie in the environmental impact evaluation for the different categories and production technologies. As LCA data must be gathered from different sources in the literature, inconsistencies in the calculations may arise. One of the difficulties experienced during the course of this work has to do with the evaluation of the space usage of the compressed hydrogen storage in the different environmental impact categories. Its inclusion might change the optimal configurations identified. Also it is observed that coal gasification is considered superior in DFF impact to steam methane reforming, concluding that these and the rest of the results presented would benefit from being more widely discussed further in the LCA expert community.

To sum up, the proposed approach enabled us to identify four redundant environmental metrics i.e., CS, ES, AE and DM), making it easier to interpret and analyze the efficient solutions to the problem. The conflicting environmental objectives (*i.e.*, CC, OLD, DM and DFF) can therefore define almost 100 % of the variance observed amongst the initial 8 environmental metrics, therefore being a good enough representative group of the bigger size problem.

Our method proves to offer valuable insights into the hydrogen SC design problem for vehicle use, suggesting process alternatives leading to significant environmental improvements and shedding light on the environmental performance of hydrogen infrastructures in different life cycle damage categories. The computational performance of the MOO-PCA methodology has proved to be computationally efficient, only implying post-optimality calculations and not adding to the complexity of the problem formulation.

It is worthwhile to make clear that the conclusions obtained about optimal hydrogen supply chain designs are proved to highly depend on the parameters, structure and objectives used for the analysis. Despite this fact, our methodology is transparent enough to analyze the main reasons for these changes and therefore generate valuable insights to help decision-making. And finally, it is general enough to be easily extended in order to handle more complex scenarios involving larger number of production, storage and transportation technologies.

Notation

Indices

B	set of chemicals indexed by b
D	set of impact categories indexed by d
I	set of hydrogen forms indexed by i
G	set of potential locations indexed by g
L	set of transportation modes indexed by l
O	set of objectives indexed by o
P	set of manufacturing technologies indexed by p
R	set of Pareto solution indexed by r
S	set of storage technologies indexed by s
T	set of time intervals indexed by t
X	set of eigenvectors
λ	set of eigenvalues indexed by e

Subsets

$IL(l)$	subset of hydrogen forms that can be transported *via* transportation mode l
$IP(p)$	subset of hydrogen forms that can be produced *via* technology p
$IS(s)$	subset of hydrogen forms that can be stored *via* technology s
$LI(i)$	subset of transportation modes for hydrogen form i
$LG(g, g')$	subset of restricted maritime links between regions g and g'
$LG'(g, g')$	subset of restricted terrestrial links between regions g and g'
$PI(i)$	subset of technologies that can produce hydrogen form i
$SI(i)$	subset of storage technologies for hydrogen form i
$SGG(g, g')$	subset of regions with active harbors that can be connected by maritime links
$SGG'(g, g')$	subset of regions with active harbors that cannot be connected by road transportation units

Parameters

av_l	availability of transportation mode l
$cc_{l,t}$	capital cost of transport mode l in period t
$cud_{l,t}$	maintenance cost of transportation mode l in period t per unit of distance traveled
$\overline{D_{g,t}}$	total demand of hydrogen in location g in period t
$distance_{g,g'}$	average distance traveled between locations g and g'
$dsat$	demand satisfaction level to be fulfilled
$fuelc_l$	fuel consumption of transportation mode l

$fuelp_{l,t}$	price of the fuel consumed by transportation mode l in period t
$ge_{l,t}$	general expenses of transportation mode l in period t
ir	interest rate
$lutime_l$	loading/unloading time of transportation mode l
$\overline{PC_p^{PL}}$	upper bound on the capacity expansion of manufacturing technology p
$\underline{PC_p^{PL}}$	lower bound on the capacity expansion of manufacturing technology p
θ	average storage period
τ	minimum desired percentage of capacity to be used
$\overline{QC_{g,g',l}}$	upper bound on the flow of materials between locations g and g' *via* transportation model l
$\underline{QC_{g,g',l}}$	lower bound on the flow of materials between locations g and g' *via* transportation model l
$\overline{SC_s^{ST}}$	upper bound on capacity expansion of storage technology s
$\underline{SC_s^{ST}}$	lower bound on capacity expansion of storage technology s
$speed_l$	average speed of transportation mode l
$tcap_l$	capacity of transport mode l
$upc_{i,g,p,t}$	value of unit production cost of hydrogen form i produced *via* technology p in location g in period t
$usc_{i,g,s,t}$	unit storage cost of hydrogen form i stored *via* technology s in location g in period t
$upcc_t$	unit transportation cost of pipelines in period t
$upoc_t$	unit operating cost of the pipelines
$usoc_t$	unit operating costs for maritime transportation and railway in period t
$wage_{l,t}$	driver wage of transportation mode l in period t
$\alpha_{g,p,t}^{PL}$	fixed investment term associated with manufacturing technology p installed in location g in period t
$\alpha_{g,s,t}^{ST}$	fixed investment term associated with storage technology s installed in location g in period t
$\beta_{g,p,t}^{PL}$	variable investment term associated with manufacturing technology p installed in location g in period t
$\beta_{g,s,t}^{ST}$	variable investment term associated with storage technology s installed in location g in period t
$\upsilon_{b,d}$	damage factor of chemical b in impact category d
ω_b^{Pr}	life cycle inventory entry of chemical b associated with hydrogen production
ω_b^{St}	life cycle inventory entry of chemical b associated with hydrogen storage
ω_b^{Tr}	life cycle inventory entry of chemical b associated with hydrogen transportation

Variables

AE	damage to ecosystem quality caused by acidification and eutrophication
$C_{g,p,t}^{PL}$	capacity of manufacturing technology p in location g

$C^{ST}_{g,s,t}$	capacity of storage technology s in location g in period t in period t
CC	damage to human health caused by climate change
$CE^{PL}_{g,p,t}$	capacity expansion of manufacturing technology p in location g in period t
$CE^{ST}_{g,s,t}$	capacity expansion of storage technology s in location g in period t
CS	damage to human health caused by carcinogenic substances
$D_{i,g,t}$	amount of hydrogen form i distributed in location g in period t
DAM_d	damage in impact category d
DFF	damage to fossil fuels
DM	damage to minerals
ES	damage to ecosystem quality caused by ecotoxic substances
FC_t	fuel cost in period t
FCC_t	facility capital cost in period t
FOC_t	facility operating cost in period t
GC_t	general cost in period t
LC_t	labor cost in period t
LCI_b	life cycle inventory of chemical b
MC_t	maintenance cost in period t
OLD	damage to human health caused by ozone layer depletion
PCC_t	pipeline capital cost in period t
POC_t	pipeline operating cost in period t
$PR_{i,g,p,t}$	production of hydrogen mode i *via* technology p in period t in location g
$Q_{i,g,g',l,t}$	flow of hydrogen mode i *via* transportation mode l between locations g and g' in period t
RE	damage to human health caused by respiratory effects
ROC_t	operating costs associated with road transportation technologies
SOC_t	ship operating cost in period t
$S_{i,g,s,t}$	amount of hydrogen in physical form i stored *via* technology s in location g in period t
SOC_t	operating cost of ships in period t
TC_t	total amount of money spent in period t
TCC_t	total transportation capital cost in period t
TDC	total discounted cost
TOC_t	transportation operating cost in period t

Integer variables

$N^{PL}_{g,p,t}$	number of plants of type p installed in location g in period t
$N^{ST}_{g,s,t}$	number of storage facilities of type s installed in location g in period t
$N^{TR}_{l,t}$	number of transportation units of type l purchased in period t

Binary variables

$Y_{g,g',l,t}$	binary variable (1 if a link between locations g and g' using transportation technology l is established, 0 otherwise)

Post-optimality analysis

CUT	threshold cut
$F_{o,r}$	value of objective o in Pareto solution r
$\overline{F_o}$	maximum value of objective o over all Pareto solutions
G_j	cumulative explained variance
$nF_{o,r}$	normalized value of objective o in Pareto solution r
j	number of the first principal components selected for analysis
x^+	the most positive element of principal component
x^-	the most negative element of principal component

CHAPTER 6

CONCLUSIONS

The important thing is not to stop questioning

Albert Einstein, 1879 – 1955

6.1 Summary

The present work presents techniques for managing uncertainty and life cycle environmental impact on two major process design problems, which are presented from the larger spatiotemporal scales involved multi-site hydrogen supply chains design to the next step of a single-site complex industrial process plant. We next present a summary of the knowledge gained during the study of each of these problems. Note that further discussions and details can be found in the results and conclusions of each corresponding chapter.

Chapter 2. Multi-objective optimization of a hydrogen supply chain – Stochastic approach

Stochastic programming coupled with risk management techniques were applied to the multi-site problem of design and planning a hydrogen supply chain in Chapter 2. This work proposes a novel decision support tool to manage the financial risk associated to the design of a hydrogen network under uncertain energy prices affecting the network operating costs *via* the worst case.

The problem addressed considers a three-echelon supply chain comprised of four different production facilities differentiated by the raw material from which hydrogen is produced (*i.e.*, natural gas, renewable energy, coal gasification and biomass), two types of storage technologies able to handle the two hydrogen product forms (*i.e.*, gas or liquid) and six several technologies for road, railway and sea transportation of the different hydrogen forms. This gives rise to a multi-scenario stochastic multi-period multi-objective MILP mathematical programming model.

The mathematical model presented is a demand driven model, which accounts for capacity expansions and economies of scale in production and storage facilities. It is important to keep in mind that the problem that arises when uncertainty in the coefficients of the objective function are considered is very different to the treatment of uncertain demands in mathematical programming frameworks.

The results presented showed a hydrogen supply chain in Spain that would benefit from having some coal gasification plants to minimize the probability of high costs in the event of undesirable high energy prices. This is a result of coal being less sensitive to market fluctuations than natural gas. Furthermore, the minimum cost solution entails a decentralized production network composed of SMR plants producing liquid hydrogen. The two-step sequential decomposition algorithm proposed to solve the model reduces in one order of magnitude the computational time of the full-space MILP, while providing solutions within 1% optimality gap.

The multi-site linear problem allows to manage the financial risk associated to energy market fluctuations, while the worst case proves to be an effective metric to control the undesirable outcomes in the uncertain parameters space.

Chapter 3. Multi-objective optimization of a hydrogen supply chain – Deterministic approach for objective reduction

In Chapter 3 we introduced a novel MOO mathematical programming framework that simultaneously optimizes the life cycle environmental performance of a hydrogen supply chain design and its economic performance. The formulation enables to evaluate independently eight independent life cycle environmental metrics that are traditionally aggregated into a single environmental impact category.

The mathematical model presented is a multi-objective multi-period MILP based on the model presented in Chapter 2. The main differences lie in the production technology dataset considered and the objective function equations. Biomass gasification for hydrogen production was removed from the analysis due to the high uncertainties associated the assessment of the environmental impact of biomass production.

Our approach, which combines MOO, LCA and PCA is comprised of two main steps that can be executed in an iterative manner.

1. The multi-objective MILP is first run and a set of Pareto fronts are generated that represent the optimal trade-off between each independent life cycle environmental impact metric and the economic performance.
2. A post-optimality analysis involving a multi-variable statistical method (*i.e.*, PCA) is then applied in order to identify redundant environmental metrics that can be removed from the analysis.

The results show that the MOO problem optimizing the cost and the environmental performance measured by eight different metric has three major different supply chain configurations. The optimum economic performance solution presents the same technologies than the ones found when optimizing the expected cost Chapter 2, while the topology of the network appears to be more centralized in the present case due to the assumption of more gradual hydrogen demand penetration. On the other hand, the minimum environmental solutions showed SC designs where SMR and WE are selected in order to reduce the environmental impact instead of coal gasification, which was shown to lead to a more robust economic performance.

The post-optimality analysis reveals that there are only four of the conflicting objectives, which would retain almost 100 % of the variance. The elimination of redundant objectives reduces the problem complexity while changing very little the Pareto structure of the problem. Among the environmental metrics considered, two distinct groups giving rise to the two different SC configurations are identified: the group of objectives that simultaneously optimize other metrics, and the group of objectives that produce diverse results in other metrics when optimized. These underlying dynamics inform about the effects that the objective choice have in the problem resolution, which might imply that for instance other environmental metrics could be negatively impacted by optimizing one of the conflicting ones, as is the case with climate change.

Chapter 4. Multi-objective optimization of industrial processes – Deterministic approach

In Chapter 4 we presented a deterministic multi-objective optimization (MOO) framework that integrates LCA principles in order to provide systematic process alternatives to improve the life cycle environmental performance of a complex industrial network.

The formulation compares three different life cycle environmental criteria against one economic performance metric.

The mathematical formulation models the operation of a complex industrial process, where several raw materials, manufacturing units and utilities are optimally combined in order to produce a set of products with predefined quality specifications. This problem gives rise to a large scale non-convex MINLP which is solved to local optimality using a customized initialization scheme and solution strategy for which computational times of less than 1 minute for normal computers are reported.

Our approach was tested using different datasets and varying demand and product quality constraints specifications, for which valuable nonintuitive solutions are obtained. For instance the model is able to identify process layouts involving the bypass or not of some units, which result in less environmental impact. Also the suitability of Pareto sets for representing important trade-offs in the process layout for the different objectives was demonstrated. Intermediate points lying between the extreme solutions corresponding to each single objective represent valuable Pareto efficient alternatives for decision makers aiming at improving the environmental performance of their processes.

The chapter closes with an initial assessment of the relevance and sensitivity of the model results in the face of different uncertainties. On the one hand parametric uncertainty is evaluated by means of two different price data sets, which lead to different economic performance but very similar environmental impact. In addition, parameter uncertainty by means of a sensitivity analysis to product quality constraints and the structural uncertainty associated to the demand satisfaction constraint flexibility are also tested. The former tests showed that the environmental performance changes very little with the limiting quality constraints, while the economic performance is more sensitive to those changes. On the other hand, the demand satisfaction appears to have an important impact on structural decisions. The methodological uncertainty associated to the performance criteria chosen for the analysis is explicitly embedded in the MOO formulation and shows different process design structures in face of parametric uncertainty.

Chapter 5. Multi-objective optimization of industrial processes – Stochastic approach

Chapter 5 presents a reformulation of the previously defined deterministic problem for the optimization of a complex single-site industrial process under uncertainty in the LCI entries. The problem is turned into a multi-scenario multi-objective stochastic model where the uncertain parameters are described through scenarios with equal probability of occurrence. The novelty of the approach is twofold: (i) it can handle any type of probability distribution of the uncertain parameters, including correlations amongst them, and (ii) for the very first time we apply the combined use of MOO and LCA to a complex single-site industrial process.

The formulation of the problem here incorporates a new environmental objective function, which is defined through a probabilistic chance constraint. In order to facilitate the resolution of the original formulation, the continuous function is discretized and approximated *via* a customized scenario generation technique for multi-variate sampling of correlated parameters, which takes into account the quality of the LCI data available. The discretized chance constraint can be calculated for a predefined confidence level that defines the minimum number of scenarios required.

Our approach overcomes the limitations associated to solving a stochastic problem that optimizes the average environmental performance, by using two stochastic metrics that allow controlling undesirable outcomes of the uncertain parameter realizations: the worst case and the downside risk. The downside risk accounts for both, the probability of exceeding an environmental limit and the deviation from such a limit, while the worst case attempts to minimize the highest impact taking place in the most unfavorable scenario. Although more simple, the worst case has the advantage of being easy and efficient to implement and computationally very advantageous.

The results of the case study presented demonstrate that the deterministic MINLP model produces solutions that can exceed the desired environmental limits when uncertainty in the LCI data is present. Our example showed that, for the problem being analyzed, the worst case and downside risk metrics produced the same results under moderate uncertainty levels. Furthermore, the minimum environmental impact solutions obtained with the stochastic approach exhibit similar structures and operational strategies than the ones presented in Chapter 4. The effect of correlations between parameters produced noticeable deviations leading to a less controllable uncertain parameter space.

6.2 List of contributions

This thesis work has advanced the state of the art of PSE mathematical programming techniques in terms of new methodological developments and applications. These include the following:

- A decision-support tool for managing financial risk in the design and planning of hydrogen supply chains under uncertainty in the operating costs was introduced. Our approach combines MOO and risk management techniques into a multi-scenario multi-period stochastic MOO MILP that evaluates all possible production, transportation and storage alternatives available for hydrogen at the moment.
- The computational complexity of this model was addressed by providing a two-step sequential algorithm that reduces the computational burden in one order of magnitude.
- An approach based on the combined use of MOO, LCA and PCA for simultaneously optimizing several life cycle environmental metrics along with the economic performance of a hydrogen supply chain was proposed. This method eliminates redundant objectives and can be seen as an alternative way to handle the methodological uncertainty associated to the selection of a specific environmental metric to be optimized.
- The combined approach of MOO and LCA was applied to a complex single-site industrial network by developing a unified systematic MINLP mathematical programming framework.
- The previous model was extended into a multi-scenario stochastic multi-objective MINLP to handle uncertainties in the LCI entries. Our approach, which can handle any type of correlated (or uncorrelated) probability distribution, allows identifying robust solutions that minimize the probability of undesirable outcomes.
- The deterministic MOO-LCA MINLP model was submitted to a set of parametric, structural and methodological tests, which in combination with its stochastic counterpart proved a useful protocol for the systematic evaluation of the parametric, structural and methodological uncertainty affecting problem formulation.
- The proposed single-site MINLP formulation for complex industrial processes was complemented by a customized solution strategy that expedites the solution strategy.
- MOO has proved to be an effective tool for incorporating environmental criteria in mathematical programming frameworks, since it allows to study inherent trade-offs arising during the analysis process alternatives.
- MOO has been shown to be an efficient tool to explore the methodological un-

certainty associated to the selection of a particular performance criteria to be optimized, since it can illustrate changes and shifts in configurations that arise from changing objectives.

- Single-site and multi-site problems may show redundant environmental objectives, and are both therefore good candidates for the application of objective reduction techniques, such as the one based on PCA.

6.3 Future work

Single-site MINLP formulation

On the modeling side:

- Study other sources of uncertainty (*i.e.*, demands, product prices, model constraints, weights, objectives, etc)
- Extend the formulation to a multi-period problem, including capital investment costs and capacity expansion of the production units.
- The substitution of unit process experimental correlations by their corresponding first principles models would be a desired final outcome to reach.
- Include more flexibility in the process design (*i.e.*, allowing different operating conditions, unit switches, etc).

On the algorithmic side:

- Several highly non-linear unit production equations can be linearized to enhance the numerical robustness of the model.
- We also intend to develop global optimization algorithms and strategies to overcome the observed problem of multi-modality. Here we may need to resort to hybrid methods that combine meta-heuristics with mathematical programming, as the use of deterministic global optimization algorithms might lead to prohibitive CPU times given the size and complexity of the model.

Multi-site MILP formulation

On the modeling side:

- Apply our stochastic framework for evaluating LCI uncertainties to the hydrogen SC covering biomass gasification and other less proved hydrogen technology options.
- Enlarge the scope of the economic performance to account for life cycle costs, although might probe challenging from the parameter gathering perspective, would allow the model to present a complete holistic economic and environmental formulation.
- Investigate further redundancies and limitations of different metrics that measure environmental impacts.

On the algorithmic side:

- Explore the use of hybrid methods to further expedite the search for Pareto optimal solutions.
- Investigate further the combination of quantitative and qualitative approaches for objective reduction and articulating decision-maker preferences.
- Study the use of multi-criteria decision-making strategies for the selection of Pareto solutions of interest.

APPENDIX A

List of publications

Research Articles

Sabio, N., Pozo, C., Guillén-Gosálbez, G., Jiménez, L., Karuppiah, R., Vasudevan, V., Sawaya, N., Farrell, J.T., 2014. Multi-objective Optimization under uncertainty of the Economic and Life Cycle Environmental Performance of Industrial Processes. *AIChE Journal* 60(6), 2098-2121

Sabio, N., Kostin, A., Guillén-Gosálbez, G., Jiménez, L., 2012. Holistic minimization of the life cycle environmental impact of hydrogen infrastructures using multi-objective optimization and principal component analysis. *International Journal of Hydrogen Energy* 37(6), 5385-5405

Sabio, N., Gadalla, M., Guillén-Gosálbez, G., Jiménez, L., 2010. Strategic planning with risk control of hydrogen supply chains for vehicle use under uncertainty in operating costs: A case study of Spain. *International Journal of Hydrogen Energy* 35(13), 6836 - 6852

Book chapters

Sabio, N., Gadalla, M., Jiménez, L., Guillén-Gosálbez G., 2010. Multi-objective optimization of a hydrogen supply chain for vehicle use including economic and financial risk metrics: A case study of Spain. *Computer Aided Chemical Engineering* 28(C), 121 - 126. DOI: 10.1016/S1570-7946(10)28021-5. ISSN: 15707946

Other publications

Sabio, N. Talking about stars, energy, optimization and science fiction, *PhDing Journal*, Chemical Engineering Department, Universitat Rovira I Virgili, 2011

Sabio, N. Let's chat about Kuhnian behaviour and scientific progress, *PhDing Journal*, Chemical Engineering Department, Universitat Rovira I Virgili, 2010

APPENDIX B

Congress contributions

Oral presentations

Sabio, N., Kostin, A., Guillén-Gosálbez, G., Jiménez, L. Optimal design and planning of sustainable hydrogen supply chains in spain: multi-objective optimization coupled with life cycle analysis and principal component analysis. *12th Mediterranean Congress of Chemical Engineering (MCCE)*, Barcelona (Spain), November 2011.

Sabio, N., Pozo, C., Guillén-Gosálbez, G., Jiménez, L., Karuppiah, R.,Vasudevan, V., Sawaya, N., Farrell, J.T. Systematic Methods for the Elimination of Redundant Life Cycle Assessment Metrics In the Multi-Objective Optimization of Industrial Processes.*American Institute of Chemical Engineers Annual Meeting*, Minneapolis (USA), October 2011.

Sabio, N., Pozo, C., Guillén-Gosálbez, G., Jiménez, L., Karuppiah, R.,Vasudevan, V., Sawaya, N., Farrell, J.T. Improving the Environmental and Economic Performance of Industrial Processes using a Multi-Objective Optimization Framework.*American Institute of Chemical Engineers (AIChE) Annual Meeting*, Salt Lake City (USA), November 2010.

Sabio, N., Gadalla, M., Jiménez, L., Guillén-Gosálbez G. Multi-objective optimization of a hydrogen supply chain for vehicle use including economic and financial risk metrics: A case study of Spain.*European symposium of computer aided process engineering (ESCAPE) 20*, Ischia (Naples), June 2010.

Sabio, N., Gadalla, M., Jiménez, L., Guillén-Gosálbez G. Risk management on the design and planning of discrete event hydrogen supply chain for vehicle use under uncertainty in production prices: A case study of Spain.*American Institute of Chemical Engineers (AIChE) Annual Meeting*, Nashville (USA), November 2009.

Sabio, N., Gadalla, M., Jiménez, L., Guillén-Gosálbez G. Strategic planning of hydrogen supply chains under uncertainty. *8th World Congress of Chemical Engineering*, Montréal (Canada), August 2009.

Sabio, N., Gadalla, M., Jiménez, L., Guillén-Gosálbez G. Multi-objective optimization of a hydrogen supply chain for vehicle use including economic and financial risk metrics: A case study of Spain. *7th PhD Poster Exhibition of URV*, Tarragona (Spain), April 2009.

.

REFERENCES

Almansoori, A., Shah, N., 2006. Design and operation of a future hydrogen supply chain: Snapshot model. *Chemical Engineering Research & Design* 84 (6), 423–438. 39, 40, 42, 46, 55, 56, 59, 60, 82, 86, 88, 100, 101

Almansoori, A., Shah, N., 2009. Design and operation of a future hydrogen supply chain: Multi-period model. *International Journal of Hydrogen Energy* 34 (19), 7883–7897. 40, 42

Althaus, H., Doka, G., Dones, R., Heck, T., Hellweg, S., et al., R. H., 2007. *Ecoinvent report no. 1– Overview and Methodology*. Tech. rep., Swiss Centre for Life Cycle Inventories. xvii, 154, 156, 192, 194, 195, 197

Amos, W., 1998. *Costs of Storing and Transporting Hydrogen*. Tech. rep., NREL – National Renewable Energy Laboratory. 56, 59, 100, 101

Antipova, E., Pozo, C., Guillén-Gosálbez, G., Boer, D., Cabeza, L. F., Jiménez, L., 2015. On the use of filters to facilitate the post-optimal analysis of the pareto solutions in multi-objective optimization. Computers and Chemical Engineering 74, 48–58. 17

Aseeri, A., Bagajewicz, M. J., 2004. New measures and procedures to manage financial risk with applications to the planning of gas commercialization in asia. *Computers & Chemical Engineering* 28 (12), 2791–2821. 178

Azapagic, A., Clift, R., 1999. The application of life cycle assessment to process optimisation. *Computers & Chemical Engineering* 23 (10), 1509–1526. 4, 14, 133

Balaman, S. Y., Selim, H., 2014. *A fuzzy multiobjective linear programming model for design and management of anaerobic digestion based bioenergy supply chains*. Energy 74, 928–940. 5

Balat, M., Kirtay, E., 2010. Major technical barriers to a "hydrogen economy". *Energy sources, Part A: Recovery utilization and environmental effects* 32 (9), 863–876. 82, 86

Ball, M., Wietschel, M., 2009a. The future of hydrogen: opportunities and challenges. *International Journal of Hydrogen Energy* 34, 615–627. 38

Ball, M., Wietschel, M., 2009b. *The Hydrogen Economy: Opportunities and Challenges*. Cambridge University Press. 38, 60

Ball, M., Wietschel, M., Rentz, O., 2007. Integration of a hydrogen economy into the german energy system: an optimising modelling approach. *International Journal of Hydrogen Energy* 32 (10–11), 1355–1368. 38

Barbaro, A. F., Bagajewicz, M. J., 2004. Managing financial risk in planning under uncertainty. *AIChE Journal* 50, 963–989. 25, 41, 52, 187

Beccali, G., Beccali, M., Cellular, M., 1997. 179

Ben-Tal, A., Nemirovski, A., 1998. Robust convex optimization. *Mathematics of Operations Research* 23 (4), 769–805. 40, 48, 185

Bernardi, A., Giarola, S., Bezzo, F., 2012. Optimizing the economics and the carbon and water footprints of bioethanol supply chains. *Biofuels, Bioproducts and Biorefining* 6 (6), 656–672. 133

Bilcke, J., Beutels, P., Brisson, M., Jit, M., 2011. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. *Medical Decision Making* 31 (4), 675–692. 6, 19, 20, 21, 22, 25

Boithias, L., Terrado, M., Corominas, L., Ziv, G., Kumar, V., Marqués, M., Schuhmacher, M., na, V. A., 2016. Analysis of the uncertainty in the monetary valuation of ecosystem services - a case study at the river basin scale. Science of the Total Environment 543, 683–690. 5

Bojarski, A., Laínez, J. M., Espuña, A., Puigjaner, L., 2009. Incorporating environmental impacts and regulations in a holistic supply chains modeling: An lca approach. *Computers & Chemical Engineering* 33 (10), 1747–1759. 4, 15, 133

Bonfill, A., Bagajewicz, M. J., Espuña, A., Puigjaner, L., 2004. Risk management in scheduling of batch plants under uncertain market demand. *Industrial and Engineering Chemistry Research* 43 (3), 741–750. 25, 52

Branke, J., Deb, K., Miettinen, K., SlowiÅĎski, R., 2012. *Multiobjective optimization: Interactive and evolutionary approaches*. Springer-Verlag, Berlin Heidelberg. 11, 15, 17

Brockhoff, D., Zitzler, E., 2009. Objective reduction in evolutionary multiobjective optimization: theory and applications. *Evolutionary Computation* 17 (2), 135–166. 17, 83

Brooke, A., Kendrik, D., Meeraus, A., Raman, R., Rosenthal, R. E., 1998. *GAMS—A User's Guide*. GAMS Development Corporation. 63, 154

Cano-Ruiz, J. A., McRae, G. J., 1998. Environmentally conscious chemical process design. *Annual Review in Energy Environment* 23, 499–536. 132

Capello, C., Hellweg, S., Badertsecher, B., Hungerbüler, K., 2005. Life-cycle inventory of waste solvent distillation: Statistical analysis of empirical data. *Environmental Science & Technology* 39 (15), 5885–5892. 180

Cattell, R. B., 2009. The scree test for the number of factors. *Multivariate Behavioral Research* 1 (2), 245–276. 97

Copado, P. J., Guillén-Gosálbez, G., Turkay, M., Jiménez, L., 2014. Milp-based decomposition algorithm for dimensionality reduction in multi-objective optimization: Application to environmental and systems biology problems. Computers and chemical engineering 67, 137–147. 17

Dantzig, G. B., 1955. *Linear programming under uncertainty.* Management Science 1, 197–206. 2

Deb, K., Saxena, D. K., 2005. *On finding Pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems.* Tech. rep., Kanpur Genetic Algorithms Laboratory, Indian Institute of Technology, Kanpur. xvi, 17, 83, 96, 97, 120, 122

DOE, 2008. *HyDRA: Hydrogen Demand and Analysis Tool.* US Department of Energy, available at:
<http://www1.eere.energy.gov/hydrogenandfuelcells/education/pdfs/hydra_joseck.pdf> [10 December 2008]. 39

DOE, 2009a. *Fuell cell technologies program – Hydrogen transition analysis workshop.* US Department of Energy, available at:
<http://www1.eere.energy.gov/hydrogenandfuelcells> [22 January 2009]. 38

DOE, 2009b. *Hydrogen Analysis Resource Center: Calculator Tools.* US Department of Energy, available at:
<http://hydrogen.pnl.gov> [28 January 2009]. 56

Dunn, S., 2002. Hydrogen futures: toward a sustainable energy system. *International Journal of Hydrogen Energy* 27, 235–264. 38

Ecoinvent Centre, 2010. *The Life Cycle Inventory Data v2. 2.* 101

EcoTransIT, 2009. Ecological transport information tool. Available at:
<http://www.ecotransit.org> [17 October 2009]. 59

Ehrgott, M., 2000a. Multicriteria optimization. Springer: Berlin. 95

Ehrgott, M., 2000b. *Multicriteria optimization.* Springer. 53, 151

El-Halwagi, M., 1997. *Pollution Prevention through Process Integration.* Academic Press, San Diego, C. A. 132

Eliceche, A. M., Corvalán, S. M., Martínez, P., 2007. Environmental life cycle impact as a tool for process optimisation of a utility plant. *Computers & Chemical Engineering*

31 (5–6), 648–656. 133

Eppen, G. D., Martin, R. K., Schrage, L., 1989. A scenario approach to capacity planning. *Operations Research* 37 (4), 517–527. 25, 187

Ferrer-Nadal, S., Guillén-Gosálbez, G., Puigjaner, L., 2008. Managing risk through a flexible recipe framework. *AIChE Journal* 54 (3), 728–740. 41

Gebreslassie, B. H., Guillén-Gosálbez, G., Boer, L. J. D., 2009a. A multi-criteria optimisation approach for the design of sustainable utility systems. *Applied Energy* 86 (9), 1712–1722. 133, 147

Gebreslassie, B. H., Guillén-Gosálbez, G., Boer, L. J. D., 2010. A systematic tool for the minimization of the life cycle impact of solar assisted absorption cooling systems. *Energy* 39 (9), 3849–3862. 133, 147

Gebreslassie, B. H., Guillén-Gosálbez, G., Jiménez, L., Boer, D., 2009b. Economic performance optimization of an absorption cooling system under uncertainty. *Applied Thermal Engineering* 29 (17–18), 3491–3500. 55

Gebreslassie, B. H., Yao, Y., You, F., 2012. *Design under uncertainty of hydrocarbon biorefinery supply chains: Multiobjective stochastic programming models, decomposition algorithm and a comparison between CVaR and downside risk*. AiChE Journal 58, 2155–2179. 5

Geisler, G., Hellweg, S., Hungerbüler, K., 2005. Uncertainty analysis in life cycle assessment (lca): Case study on plant-protection products and implications for decision making. *nternational Journal of Life Cycle Assessment* 10 (3), 184–192. 180

Geletu, A., Li, P., 2002. Recent developments in computational approaches to optimization under uncertainty and application in process systems engineering. *ChemBioEng reviews* 1 (4), 170–190. 5

Goel, T., Vaidyanathan, R., Haftka, R. T., Shyy, W., Queipo, N. V., Tucker, K., 2007. Response surface approximation of pareto optimal front in multi-objective optimization. *Computer Methods in Applied Mechanics and Engineering* 196 (4–6), 879–893. 97

Grossmann, I. E., 1989. *MINLP optimization strategies and algorithms for process synthesis*. Tech. rep., Chemical Engineering Department, Carnegie Mellon. 9, 22

Grossmann, I. E., 2002. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and Engineering* 3, 227–252. 4, 9, 153

Grossmann, I. E., 2004. Challenges in the new millennium: product discovery and design, enterprise and supply chain optimization,global life cycle assessment. *Computers and Chemical Engineering* 29, 29–39. 6, 9, 40

Grossmann, I. E., 2005. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE Journal* 51, 1846–1857. 6, 9, 40

Grossmann, I. E., 2012. Advances in mathematical programming models for enterprise-wide-optimization. *Computers and chemical engineering* 76, 2–18. 6, 9, 27

Grossmann, I. E., 2014. Challenges in the application of mathematical programming in the enterprise-wide optimization of process industries. Theoretical foundations of chemical engineering 48 (5), 555–573. 9, 11, 21, 24, 25, 27, 29

Grossmann, I. E., Drabbant, R., Jain, R. K., 1982. Incorporating toxicology in the synthesis of industrial chemical complexes. *Chemical engineering communications* 17 (1-6), 151–170. 4

Grossmann, I. E., Guillén-Gosálbez, G., 2010. Scope for the application of mathematical programming techniques in the synthesis and planning of sustainable processes. *Computers & Chemical Engineering* 34 (9), 1365–1376. xiii, 1, 4, 8, 9, 13, 132, 133

Grossmann, I. E., Westerberg, A. W., 2000. Challenges in process systems engineering. *Aiche Journal* 46 (9), 1700–1703. 2, 6, 8, 9

Guillén-Gosálbez, G., 2011a. A novel milp-based objective reduction method for multi-objective optimization: Application to environmental problems. *Computers & Chemical Engineering* 35 (8), 1469–1477. 17

Guillén-Gosálbez, G., 2011b. A novel milp-based objective reduction method for multi-objective optimization: application to environmental problems. *Computers & Chemical Engineering* 35 (8), 1469–1477. 83

Guillén-Gosálbez, G., Badell, M., Puigjaner, L., 2007. A holistic framework for short-term supply chain management integrating production and corporate financial planning. *International Journal of Production Economics* 106 (1), 288–306. 90

Guillén-Gosálbez, G., Caballero, J. A., Jiménez, L., 2008. Application of life cycle assessment to the structural optimization of process flowsheets. *Industrial & Engineering Chemistry Research* 47 (3), 777–789. 4, 15, 133, 147

Guillén-Gosálbez, G., Espuña, A., Puigjaner, L., 2006a. Addressing the scheduling of chemical supply chains under demand uncertainty. *AIChE Journal* 52 (11), 3864–3881. 40

Guillén-Gosálbez, G., Grossmann, I. E., 2009. Optimal design and planning of sustainable chemical supply chains under uncertainty. *AIChE Journal* 55 (1), 99–121. 5, 48, 180, 202

Guillén-Gosálbez, G., Grossmann, I. E., 2010. A global optimization strategy for the environmentally conscious design of chemical supply chains under uncertainty in the damage assessment model. *Computers & Chemical Engineering* 34 (1), 42–58. 5, 37, 40, 42, 133, 180, 202

Guillén-Gosálbez, G., Mele, F. D., Bagajewicz, M. J., Espuña, A., Puigjaner, L., 2005. Multiobjective supply chain design under uncertainty. *Chemical Engineering Science*

60 (6), 1535–1553. 40, 41

Guillén-Gosálbez, G., Mele, F. D., Espuña, A., Puigjaner, L., 2006b. Addressing the design of chemical supply chains under demand uncertainty. *Industrial & Engineering Chemistry Research* 45 (22), 7566–7581. 40

Guillén-Gosálbez, G., Mele, F. D., Grossmann, I. E., 2010. A bi-criterion optimization approach for the design and planning of hydrogen supply chains for vehicle use. *AIChE Journal* 56 (3), 650–667. 4, 5, 15, 37, 39, 42, 53, 83, 86, 133, 147

Guillén-Gosálbez, G., Puigjaner, L., 2008. Towards an integrated framework for supply chain management in the batch chemical process industry. *Computers & Chemical Engineering* 32 (4–5), 650–670. 133, 147

Guinée, J. B., Gorrée, M., Heijungs, R., Huppes, G., Kleijn, R., de Koning et al., A., 2002. *Handbook on life cycle assessment. Operational guide to the ISO standards.* Kluwer Academic Publishers, Dordrecht, NL. 4, 13, 133, 147

Gutiérrez, E., Lozano, S., Moreira, M. T., Feijoo, G., 2010. Assessing relationships among life-cycle environmental impacts with dimension reduction techniques. *Journal of Environmental Management* 91 (4), 1002–1011. 95

Guttman, L., 1954. Some necessary conditions for common factor analysis. *Psychometrika* 19 (2), 149–162. 96

Haimes, Y., Lasdon, L., Wismer, D., 1971. On a bi-criterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man and Cybernetics* 1 (3), 296–297. 12, 151, 194

Hajimiragha, A., Fowler, M. W., Ca'nizares, C. A., 2009. Hydrogen economy transition in ontario canada considering the electricity grid constraints. *International Journal of Hydrogen Energy* 34 (13), 5275–5293. 38

Hoffman, L., Weidema, W. P., Kristiansen, K., Ersboll, A. K., 1995. *Statistical analysis and uncertainties in relation to LCA. Special reports no. 1.* Tech. rep., *Nordic Council of Ministers*, Copenhaguen, Denmark. 179

Hugo, A., Pistikopoulos, E. N., 2005. Environmentally conscious long-range planning and design of supply chain networks. *Journal of Cleaner Production* 13 (15), 1471–1491. 4, 15, 133

Hugo, A., Rutter, P., Pistikopoulos, S., Amorelli, A., Zoia, G., 2005. Hydrogen infrastructure strategic planning using multi-objective optimization. *International Journal of Hydrogen Energy* 30 (15), 1523–1534. 39, 55, 83, 133

Huijbregts, M. A. J., 1998. Part i: A general framework for the analysis of uncertainty and variability in life cycle assessment. *International Journal of Life Cycle Assessment* 3 (5), 273–280. 5, 179

Huijbregts, M. A. J., 2001. *Uncertainty and variability in environmental life-cycle as-*

sessment. PhD Thesis, Universiteit van Amsterdam. 5, 6, 20, 25

Hung, M. L., Ma, H. W., 2009. Quantifying system uncertainty of life cycle assessment based on monte carlo simulation. *International Journal of Life Cycle Assessment* 14 (1), 19–27. 180

HyWays, 2004. *HyWays scoping report–Hydrogen penetration rates for stationary and mobile applications*. Tech. rep., European Commission. 98

HyWays, 2010. *Análisis del mapa de ruta del H_2 en España*. Plataforma tecnológica Española del hidrógeno y de las pilas de combustible – Grupo de Análisis de Capacidades. 98

IEA, 2015. *World Energy Outlook 2015*. Tech. rep., OECD Publishing, Paris. 1

IMF, 2009. International monetary fund. Available at: <http://www.imf.org> [24 September 2009]. 56, 101

INE, 2009. Instituto nacional de esdatística (Spanish statistical office). Available at: <http://www.ine.es> [15 September 2009]. 56, 60, 101

Ingason, H., Ingolfsson, H., Jensson, P., 2008. Optimizing site selection for hydrogen production in iceland. *International Journal of Hydrogen Energy* 33 (14), 3632–3643. 39, 82

ISO, 1997–2000. Series of standards on environmental management and life-cycle assessment: ISO 14040, 14041, 14042, 14043. 94

Iyer, R. R., Grossmann, I. E., 1998. A bilevel decomposition algorithm for long-range planning of process networks. *Industrial & Engineering Chemistry Research* 37 (2), 474–481. 53

Jackson, D. A., 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology* 74 (8), 2204–2214. 97

Jackson, J. E., 2003. *A user's guide to principal components*. Wiley-Interscience. 83, 95

Janak, S. L., Lin, X., Floudas, C. A., 2007. A new robust optimization approach for scheduling under uncertainty - ii. uncertainty with known probability distribution. *Computers & Chemical Engineering* 31 (3), 171–195. 48, 186

Kaiser, H. F., 1960. The application of electronic computer to factor analysis. *Educational & Psychological Measurement* 20 (1), 141–151. 96

Kamarudin, S., Daud, W., Yaakubb, Z., Misron, W., Anuar, N. Y., 2009. Synthesis and optimization of future hydrogen energy infrastructure planning in peninsular malaysia. *International Journal of Hydrogen Energy* 34 (5), 2077–2088. 40

Karuppiah, R., Grossmann, I. E., 2012. A langrangean-based branch-and-cut algorithm for global optimization of non-convex mixed integer non-linear programming problems with decomposable structures. *Journal of Global optimization* 41 (2), 163–186. 4

Kelly, N., Gibson, T., Cai, M., Spearot, J., Ouwerkerk, D., 2010. Development of a renewable hydrogen economy: Optimization of existing technologies. *International Journal of Hydrogen Energy* 35 (3), 892–899. 38

Kennedy, D. J., Montgomery, D. C., Quay, B. H., 1996. Data quality - stochastic environmental life cycle assessment modeling: A probabilistic approach to incorporating variable input data quality. *International Journal of Life Cycle Assessment* 1 (4), 199–207. 179

Kim, J., Lee, Y., Moon, I., 2008. Optimization of a hydrogen supply chain under demand uncertainty. *International Journal of Hydrogen Energy* 33 (18), 4715–4729. 39, 42, 55, 59, 82, 100

Kim, J., Moon, I., 2008a. The role of hydrogen in the road transportation sector for a sustainable energy system: A case study of korea. *International Journal of Hydrogen Energy* 33 (24), 7326–7337. 38

Kim, J., Moon, I., 2008b. Strategic design of hydrogen infrastructure considering cost and safety using multiobjective optimization. *International Journal of Hydrogen Energy* 33 (21), 5887–5896. 39, 42, 55

Konieczny, A., Mondal, K., P., T. W., Dydo, 2008. Catalyst development for thermocatalytic decomposition of methane to hydrogen. *International Journal of Hydrogen Energy* 33 (1), 264–272. 86

Koroneos, C., Dompros, A., Roumbas, G., Moussiopoulos, N., 2004. Life cycle assessment of hydrogen fuel production processes. *International Journal of Hydrogen Energy* 29 (14), 1443–1450. 82

Koroneos, C., Dompros, A., Roumbas, G., Moussiopoulos, N., 2005. 82, 86, 101

Kostin, A., Guillén-Gosálbez, G., Turkay, M., Jiménez, L., 2015. Dimensionality reduction applied to the simultaneous optimization of the economic and life cycle environmental performance of supply chains. International Journal of production economics 159, 223–232. 17

Krzanowski, W. J., 2000. *Principles of multivariate analysis: a user's perspective.* Oxford University Press. 95

Laínez, J. M., Guillén-Gosálbez, G., Badell, M., Espuña, A., Puigjaner, L., 2007. Enhancing corporate value in the optimal design of chemical supply chains. *Industrial & Engineering Chemistry Research* 46 (23), 7739–7757. 90

Law, A. M., Kelton, W. D., 2000. *Simulation Modeling & Analysis*, 3rd Edition. McGraw-Hill, New York. 60, 193, 197

Li, Z., Gao, D., Chang, L., Liu, P., Pistikopoulos, E. N., 2008. Hydrogen infrastructure design and optimization: A case study of china. *International Journal of Hydrogen*

Energy 33 (20), 5275–5286. 39, 55, 83

Lin, Z., Chen, C., Ogden, J., Fan, Y., 2008a. The least-cost hydrogen for southern california. *International Journal of Hydrogen Energy* 33 (12), 3009–3014. 39, 82

Lin, Z., Ogden, J., Fan, Y., Chen, C., 2008b. The fuel-travel-back approach to hydrogen station siting. *International Journal of Hydrogen Energy* 33 (12), 3096–3101. 39

Linderoth, J., Shapiro, A., Wright, S., 2006. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research* 142 (1), 215–241. 40

Linnhoff, B., 1993. Pinch analysis–a state-of-the-art overview. *Chemical Engineering Research & Design* 71 (5), 503–522. 132

Liu, P., Pistikopoulos, E. N., Li, Z., 2010. A multi-objective optimization approach to polygeneration energy systems design. *AIChE Journal* 56 (5), 1218–1234. 133

Marler, R. T., Arora, J. S., 2004. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization* 26 (6), 369–395. 95

McCabe, W., Smith, 2001. *Unit Operations of Chemical Engineering*. McGraw-Hill Book Company Inc., New Yark. 22

Mele, F. D., Espuña, A., Puigjaner, L., 2005. 133, 147

Mele, F. D., Kostin, A., Guillén-Gosálbez, G., Jimenez, L., 2011. Multiobjective model for more sustainable fuel supply chains. a case study of the sugar cane industry in argentina. *Industrial & Engineering Chemistry Research* 50 (9), 4939–4958. 4, 5, 133, 147

Melo, M., Níkel, S., da Gama, F. S., 2009. Facility location and supply chain management–a review. *European Journal of Operation Research* 196 (2), 401–412. 40

MITYC, 2008. Spanish ministry of industry, tourism and commerce: Reports and statistics. Available at:
<http://www.mityc.es> [10 December 2008]. 56

NRC, 2004. *The hydrogen economy: Opportunities, Costs, Barriers and R&D needs.* Tech. rep., National Research Council and Committee on Alternatives and Strategies for Future Hydrogen Production and Use and National Academy of Engineering. National Academy of Sciences U.S.,National Academies Press Edition. 38

NRC, 2009. Nacional research council. Available at:
<http://www.nrc.gov> [15 January 2009]. 60

Oliva, D. G., Guillén-Gosálbez, G., Mateo-Sanz, J. M., Jiménez-Esteller, L., 2013. Milp-based clustering method for multi-objective optimization: Application to environmental problems. Computers and Chemical Engineering 56, 202–217. 17

OSE, 2009. Sustainability observatory of spain: Energy consumption by sectors. Available at:
<http://www.sostenibilidad-es.org> [20 January 2009]. 56

Osmani, A., Zhang, J., 2014. *Economic and environmental optimisation of a large scale sustainable dual feedstock lignocellulosic-based bioethanol supply chain in a stochastic environment.* Applied Energy 114, 572–587. 5

Padró, C. E. G., Putsche, V., 1999. *Survey of the Economics of Hydrogen Technologies.* Tech. rep., National Renewable Energy Laboratory. 86, 100

Papandreou, V., Shang, Z., 2008. A multi-criteria optimisation approach for the design of sustainable utility systems. *Computers & Chemical Engineering* 32 (7), 1589–1602. 133

Petersen, E. H., 1997. Life-cycle assessment of building components. handling uncertainties in lcas. 179

Pieragostini, C., Mussati, M. C., Aguirre, P., 2012. On process optimization considering lca methodology. *Journal of Environmental Management* 96 (1), 43–54. 4

Pistikopoulos, E. N., Stefanis, S. K., 1998. Optimal solvent design for environmental impact minimization. *Computers & Chemical Engineering* 22 (6), 717–733. 133

Pozo, C., Ruíz-Femenia, R., Caballero, J. A., Guillén-Gosálbez, G., Jiménez, L., 2012. On the use of principal component analysis for reducing the number of environmental objectives in multi-objective optimization: Application to the design of chemical supply chains. *Chemical Engineering Science* 69 (1), 146–158. 17, 147

Reap, J., Roman, F., Duncan, S., Bras, B., 2008. A survey of unresolved problems in life cycle assessment. part 2: Impact assessment and interpretation. *International Journal of Life Cycle Assessment* 13 (5), 374–388. 179

Reyes-Labarta, J. A., Salcedo-Díaz, R., Ruiz-Femenia, R., Guillén-Gosálbez, G., Caballero, J. A., 2014. *Handling of uncertainty in life cycle inventory by correlated multivariate lognormal distributions: Application to the design of supply chain networks.* Computer Aided Chemical Engineering 33, 1075–1080. 5

Rosenthal, R., 2008. GAMS — A user's guide. GAMS Development Corporation: Washington. 104

Ruíz-Femenia, R., Guillén-Gosálbez, G., Jiménez, L., Caballero, J. A., 2013. Multi-objective optimization of environmentally conscious chemical supply chains under demand uncertainty. *Chemical Engineering Science* 95, 1–11. 133, 180

Sabio, N., Gadalla, M., Guillén-Gosálbez, G., Jimenez, L., 2010. Strategic planning with risk control of hydrogen supply chains for vehicle use under uncertainty in operating costs: A case study of spain. *International Journal of Hydrogen Energy* 35 (13), 6836–6852. 29, 38, 83, 86, 133

Sabio, N., Kostin, A., Guillén-Gosálbez, G., Jiménez, L., 2012. Holistic minimization of the life cycle environmental impact of hydrogen infrastructures using multi-objective optimization and principal component analysis. *International Journal of Hydrogen Energy* 37 (6), 5385–5405. 17, 30, 82

Sabio, N., Pozo, C., Guillén-Gosálbez, G., Jiménez, L., Karuppiah, R., Vasudevan, V., Sawaya, N. W., Farrell, J. T., 2014. Multiobjective optimization under uncertainty of the economic and life-cycle environmental performance of industrial processes. *AIChE Journal* 60 (6), 2098–2121. 5, 7, 31, 178, 184, 188, 193

Sahinidis, N., 2004. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering* 28 (6–7), 971–983. 2, 5, 20, 40, 178, 184

Salcedo, R., Antipova, E., Boer, D., Jiménez, L., Guillén-Gosálbez, 2012. Multi-objective optimization of solar rankine cycles coupled with reverse osmosis desalination considering economic and life cycle environmental concerns. *Desalination* 286, 358–371. 5

Sargent, R., 2005. Process systems engineering: a retrospective view with questions for the future. *Computers and chemical engineering* 29, 1237–1241. 22

Sawaya, N., 2006. *TReformulations, relaxations and cutting planes for generalized disjunctive programming*. PhD Thesis, Chemical Engineering Department, Carnegie Mellon University. 4

Seydel, P., Wietschel, M., 2007. *Report on the HyWays scenario assumptions*. Tech. rep., Fraunhofer Institute for Systems and Innovation Research. 98

Shapiro, J., 2004. Challenges of strategic supply chain planning and modelling. *Computers & Chemical Engineering* 28 (6–7), 855–861. 40

Simchi-Levi, D., Kamisky, P., Simchi-Levi, E., 2000. Designing and managing the supply chain. Concepts, strategies, and case studies. Irwin McGraw-Hill. 9, 88

Spath, P. L., Mann, M. K., 2001. Life cycle assessment of hydrogen production via natural gas steam reforming. Tech. rep., NREL–National Renewable Energy Laboratory. 82, 86, 101

Spath, P. L., Mann, M. K., 2004. Life cycle assessment of renewable hydrogen production via wind/electrolysis. Tech. rep., NREL–National Renewable Energy Laboratory. 82, 101

Stefanis, S. K., Livingston, A. G., Pistikopoulos, E. N., 1997. Environmental impact considerations in the optimal design and scheduling of batch processes. *Computers & Chemical Engineering* 21 (10), 1073–1094. 133

Sugiyama, H., Fukushima, Y., Hirao, M., Hellweg, S., Hungerbüler, K., 2005. Using standard statistics to consider uncertainty in industry-based life cycle inventory databases. *nternational Journal of Life Cycle Assessment* 10 (6), 399–405. 180

Téno, J. L. C. J. F. L., 1996. Life cycle analysis with ill-defined data and its application to building products. *International Journal of Life Cycle Assessment* 1 (2), 90–96. 179

Transmar S. L., 2009. *Transmar ship & forwarding agency*. Available at: <http://www.transmar.net/index.html> [18 February 2009]. 56, 59, 101

Van den Heever, S. A., Grossmann, I. E., 2003. A strategy for the integration of production planning and reactive scheduling in the optimization of a hydrogen supply network. *Computers & Chemical Engineering* 27 (12), 1813–1839. 39

Vaskan, P., Guillén-Gosálbez, G., Turkay, M., Jiménez, L., 2014. Multiobjective optimization of utility plants under several environmental indicators using an milp-based dimensionality reduction approach. Industrial and Engineering Chemistry Research 53 (50), 19559–19572. 17

Wang, Y., Smith, R., 1995. Wastewater minimization with flowrate constraints. *Chemical Engineering Research & Design* 73 (8), 889–904. 132

Weidema, B. P., Wesnaes, M. S., 1996. Data quality management for life cycle inventories - an example of using data quality indicators. *Journal of Cleaner Production* 4, 167–174. 188, 190, 192, 195

Weissman, W., Farrell, J., Sasaki, S., Akihama, K., Nakakita, K., 2012. Diesel fuel formulation for reduced emissions. CA Patent 2,478,119. 38

Weissman, W., Farrell, J., Schilowtiz, A., Johnson, J., Ueda, T., Akihama, K., 2011. Tuning fuel composition for driving cycle conditions in spark ignition engines. CA Patent 2,439,154. 38

World Resources Institute, 2005. Earthtrends data tables: Climate and atmosphere. Available at: <http://earthtrends.wri.org/pdf_library/data_tables/cli2_2005.pdf> [20 February 2009]. 38

WRI, 2005. Earthtrends data tables: Energy consumption by sector – world resources institute. Available at: <http://earthtrends.wri.org/pdf_library/data_tables/ene3_2005.pdf> [20 February 2009]. 38

Yang, C., Ogden, J., 2007. Determining the lowest-cost hydrogen delivery mode. *International Journal of Hydrogen Energy* 32 (20), 268–286. 59

Zamboni, A., Shah, N., Bezzo, F., 2009. Spatially explicit static model for the strategic design of future bioethanol production systems. 2. multi-objective environmental optimization. Energy & Fuels 23, 5134–5143. 5

Zelm, R. V., Huijbregts, M. A. J., Posthuma, L., Wintersen, A., Meent, D. V. D., 2009. Pesticide ecotoxicological effect factors and their uncertainties for freshwater ecosystems. *International Journal of Life Cycle Assessment* 14 (1), 43–51. 5, 180

Zuur, A., Ieno, E., Smith, G., 2007. *Analysing ecological data*. Springer. 95

Note: The numbers listed behind the full stop of each bibliography entry refer to the pages on which the entry is cited.

This thesis presents a general mathematical programming framework for the design of more sustainable energy systems and process industries under uncertainty. The approach presented consists on the formulation of a backbone multi-objective model, which is initially tested in a classical deterministic form, to then be reformulated as a stochastic programming problem. The first half of the work is devoted to the design of robust hydrogen supply chains under uncertain market conditions and different environmental metrics. The second half of the thesis brings life cycle assessment into a novel industrial-scale nonlinear multi-objective optimization framework for systematically designing robust process plants. Risk management metrics are used for controlling the impact of the uncertainty associated to the economic and environmental performance indicators. Principal component analysis is shown to effectively reduce the dimensionality of problems with a large number of objectives to smaller non-redundant and representative sets. Life cycle environmental impact indicators are proved as efficient additional optimization criteria capable of uncovering robust and more sustainable process systems designs. Finally uncertainty is analysed using a multi-dimensional perspective that allows to represent complex systems-based research problems.

Universitat Rovira i Virgili