

École polytechnique de Louvain

Optimal strategic deployment of green hydrogen production and supply infrastructures in Belgium based on MINLP model, solved with multi-variate metaheuristic algorithm

Author: **Matthias GEURTS**
Supervisor: **Francesco CONTINO**
Readers: **Davide TONELLI, Philippe CHEVALIER**
Academic year 2021–2022
Master [120] in Mathematical Engineering

Abstract

The search for energy has always been at the core of human existence. Nowadays, Humanity has finally realized that continuing to burn fossil fuels and release billions of tons of CO_2 in the atmosphere year after year is not an option. Climate change is clearly visible, with its consequences: temperatures increasing, polar ice melting, meteorological disasters, impact on living species and populations across the globe.

In response, ambitious targets have been set to bring the world to net zero emissions in the coming decades. The replacement of today's energy sources with renewable, stable and cheap sources is key, however it represents a challenge from a societal, economical and technological point of view. Among the solutions to be adopted, the introduction of a new energy carrier like hydrogen can fasten the defossilization of sectors heavily dependent on fossil sources, for which few mature technological alternatives exist.

This master thesis aims at developing an approach to optimize the future hydrogen production-distribution system of a country, while taking into account major sources of uncertainty. Indeed, integrating unknown future hydrogen demand and energy costs changes the way the design of the system balances imports and local production. That is what we will show by applying our developments to Belgium as case study.

We develop an algorithm combining mixed-integer non-linear programming with a multi-stage development of scenarios to integrate sources of uncertainty. More specifically, we develop a two-level algorithm where an Ant-Colony metaheuristic embeds a classic solver, which allows to clearly separate linear from non-linear objectives and constraints.

Our results are promising, as seen on the application to the case study. Our approach finds stable network configurations at different points in time (2030, 2040, 2050), at minimum capex and opex levels. The computational time remains reasonable for different problem sizes and constraint types, in the order of a couple of minutes to find optimal solutions. Our approach is generic enough to introduce new dimensions to the objective function, new constraints, more time-steps and new technology of production and transport.

Based on our work, we see a couple of dimensions to explore in order to bring our development to a more robust, or even industrial level. In particular, we believe that introducing new choices for the production and transport technologies, testing with smaller time-steps and introducing new uncertain variables in sets of plausible scenarios should allow the model and algorithm to determine more realistic, economically-viable solutions while keeping options open for new technology developments in the future.

Contents

1	Introduction	4
1.1	Hydrogen: Possibilities and Impact	4
1.2	Problem Definition	7
1.3	Overview of the Master Thesis	8
1.4	Acknowledgements	9
2	Stochastic Model: Theoretical Review	10
2.1	Hydrogen Supply Chain (HSC)	10
2.2	Two foundational papers in stochastic programming for HSC	12
2.3	Approaches to Uncertainty in the Literature	17
2.4	Multi-Stage Modeling	18
3	Model	22
3.1	Key Modeling Choices	22
3.2	Objective Functions	23
3.3	Constraints	24
3.4	Formal Declarations	28
4	Belgian Case Study: Key Dimensions and Design Choices	35
4.1	Discrete Time Evolution	35
4.2	Topography as a Network	35
4.3	Hydrogen Demand	37
4.4	Hydrogen Production	38
4.5	Hydrogen Transport	39
4.6	Electricity Market	41
4.7	Hydrogen Import	42
4.8	Trees of Scenarios - Uncertainty Combinations	42
4.9	Implementation Choices	44
5	Results	47
5.1	Experimentation Plan and Graphical Representation of Solutions .	47

5.2	Base Case: Unconstrained Capacity and Utilization	50
5.3	Variation 1: Capacity Concentration and Utilization Factor	51
5.4	Variation 2: Uncertainty in Cost of Electricity Import	56
5.5	Variation 3: Biased Probability Distribution	57
5.6	Variation 4 : Large Variation in Hydrogen Import Costs	58
5.7	Complexity Analysis	60
5.8	Assessment of Our Approach and Results	65
6	Conclusion	67

Chapter 1

Introduction

Net zero is certainly one of the most difficult and important challenges humanity will have to tackle over the next decades if we want to preserve the Earth as our home. Mastering green sources of energy at scale will be key to get there and hydrogen is part of the tools that could result to be fundamental.

This master thesis aims to contribute to this objective by developing a new approach to optimize the future deployment of the infrastructure for the production and distribution of hydrogen in Belgium, taking into account different sources of non-linearity and uncertainty.

This introductory chapter successively presents the broad context of hydrogen and its of application, introduces the specific problem and the solution we have developed, and gives an overview of this report.

1.1 Hydrogen: Possibilities and Impact

The impact of human societies on the planet has become a major subject of interest over the past decades. With the Paris Agreement on Climate Change, the international community agreed to reduce greenhouse gas emissions and contain the average increase of temperature to 2°C above pre-industrial levels. This objective, if sometimes not considered ambitious enough, will still require significant efforts from all countries in the coming years. Likewise, Belgium targets a reduction of 80% of its emissions by 2050. In the fight against climate change, people also put a lot of hope in new technologies. In particular, hydrogen comes as a solution for the decarbonization of a number of major activity sectors. The Hydrogen Import coalition predicts an increasing role of hydrogen to support Europe's energy transition [9].

Hydrogen is a synthetic fuel that can be produced by diverse methods and has potential of application among a large number of end-uses: : transport, chemicals, steel, power and heat generation [24]. There is yet an important difference to make between different production options. The so-called "green hydrogen" refers to a low carbon production based on renewable energies. The "grey hydrogen" is produced using gas and thus implying an important environmental impact. Finally, the "blue hydrogen" option uses natural gas, but also new technologies for the capture of CO_2 emissions. A green hydrogen economy will necessarily require an abundant availability of renewable sources for the production of the carrier.

Unfortunately, not all countries have that much green energy available that they could use for the production of hydrogen. Certain regions in the world are, as one could say, "blessed" with (stable and cheap) solar and wind energy potential - two of the most regarded options for green energy production. The next steps would then be to transport the newly produced energy to regions that cannot leverage these renewable sources but this would come at great scale and cost. Indeed, the most industrialized regions of central Europe are often also the ones with the least potential [17]. Hydrogen is a solution to this problem. Its versatility is a main advantage. It can be transported the same way as traditional fossil fuels. In fact, a portion of the already existing infrastructures could be adapted to transport this new carrier at a reasonable cost.

A second problem often addressed of the green energy production, is its unpredictability. Of course a solar panel only produces when there is sun. Meteorological conditions greatly influence wind turbines production. However, the energy demand from the network does not follow the same variations. It has become necessary to store energy and be able to retrieve it with minimum losses. If battery capacities have massively increased in the last years, the molecule storage option still seems more attractive. Battery technologies remain expensive and require scarce raw materials for production [9]. Once converted into chemical bonds between molecules, the energy can be stored for longer time and be used when needed.

Nevertheless, using hydrogen presents certain challenges. The fuel needs to be produced using other forms of energy, it is not a direct replacement of traditional natural gas or oil that can be pumped from the ground. It also has a lower volumetric energy density than traditional fuels and requires certain temperature and pressure conditions to be transported and stored, which inevitably increases the cost and complexity of the infrastructure. A lot of the technologies are not at maturity yet and they are not ready to be used at scale. But it is also encouraging to see

that improvement in costs and usability of hydrogen technologies will continue to accelerate with penetration as learning curves from other technologies have shown in the last decades (e.g., microelectronics, solar panels, batteries, etc).

The green hydrogen market is still very new, and uncertain. In the favorable scenarios, if all actors played their role, hydrogen could become economically competitive by 2030 and the "market for low-carbon hydrogen and associated synthetic fuels could reach \$1 trillion by the middle of this century" [13]. The report [24] decomposes the so-called "hydrogen economy" in four segments: production, distribution, conversion, and applications. Our work will focus on the production and distribution steps of the problem.

The first strategy would be for European countries to produce their own hydrogen directly with locally produced green electricity. There are also several regions that present the right conditions to become net hydrogen exporters in the next decades. Those countries offer political stability and green energy potential. The hydrogen could be exported to our regions by sea, following the navigation routes presented in figure 1.1. The best strategies will leverage these two approaches in a balanced, optimal way.



Figure 1.1: Hydrogen Import Coalition [9] : shipping routes prediction for 2050

1.2 Problem Definition

As we have seen, a number of technical challenges will have to be overcome to make the hydrogen economy become a reality. Countries like Belgium will not become "green hydrogen champions" because of lack of renewable power sources. However, our existing gas infrastructure combined with new production and distribution capacities, will constitute a practical solution to achieve our decarbonization ambition.

In this context, this master thesis focuses on the production-distribution infrastructure. The algorithm developed provides an optimal design for the network based on a cost minimization. Our problem is of the type of a supply problem, where first, decisions are made to design a future supply system, and then, there is a second round of decisions for the allocation of the resources to satisfy a given demand.

We model the spatial complexity of our problem with a network. The space is decomposed in regions represented by nodes. Each node gathers the demand, carrier availability and production potential of one region. We are interested to build the production capacities but also transport infrastructures between regions, hence building the entire network.

We are here interested in green hydrogen only. This means that we focus on hydrogen produced using green electricity. This implies a focus on electrolysis technologies – currently the most viable for large-scale deployment [26].

We want to solve the problem of the design of the infrastructures, in order to satisfy a future uncertain market that is still in development. The two sources of uncertainty that will be integrated in our approach are hydrogen demand and import prices of materials. In our case, those are mainly hydrogen and electricity. The type of uncertainty that we are working with on this case is thus related to long-term evolution of the model parameters and not short-term demand forecasting and full-filling.

To draft the evolution of the network, we will use a model with successive time steps of approximately 10 years each. As opposed to an infinite process, the model uses a finite horizon with a determined set of time-steps. At each time step, new information is added about real demand and cost evolution up to that point. This model represents this progressive arrival of information. Using this new information, the solution gives us new decisions to make to upgrade the network design for the next round. Part of those decisions will only come to operation at the next time-step – this is where design under uncertainty takes all its meaning.

The model we are working on is of "Mixed-Integer Non-Linear Programming" (MINLP) type. The important boolean decisions, such as deciding to build more capacity, are represented as integer variables. Moreover, the model can handle some non-linearity like the installation CAPEX for new hydrogen production capacities. This non-linearity is very often directly linearized in literature. Actually, the problem contains other sources of uncertainty, such as pressure differences between the different parts of the system for example. Whereas this is not the subject of this master thesis, our model was developed versatile enough to support such new features. Moreover, one of the objectives of the thesis was to add uncertainty in an already existing model to contribute to the development of an algorithmic tool to study chemical supply chain design. The metaheuristic used to handle the uncertainty is the Ant-Colony [20]. Our solution had to integrate our new algorithm in this already existing metaheuristic.

1.3 Overview of the Master Thesis

The rest of this document is structured in 5 successive chapters following this introduction.

Chapter 2 briefly presents our review of the literature linked to the problem at hand with a focus on stochastic theory. In particular, we focus on key constraints and building blocks of our problem, namely: how uncertainty is introduced in different models, how authors leverage multi-stage approaches, and finally key (meta-)heuristics used to develop robust solutions.

Chapter 3 gives a formal definition of the solution model of our problem. We successively detail out the objective function we are working on, translate technical and physical constraints mathematically, and finally provide all necessary formal definitions of variables, parameters, functions.

Chapter 4 describes our case study, namely the hydrogen supply chain design in Belgium. It covers all building blocks of the "real" problem to solve, such as location of existing and new assets, production and demand evolution forecasts, production and transport costs, hydrogen import sources. It finally provides a brief description of our key implementation design choices, in particular our two-layer algorithm embedding linear optimization within an Ant-Colony metaheuristic.

Chapter 5 is where we present and discuss the results achieved in this work. We successively review versions of our model and solution, with different sets of parameters and constraints. We also address the complexity of our solution, i.e. the computation time taken to tackle problems of increasing size. Finally we discuss dimensions we could add to our model to find optimal solutions to more complex (real) problems.

Chapter 6 finally concludes the work with a recap of our results, a critical review of the approach and suggestions for further work.

1.4 Acknowledgements

The work presented here has been influenced by a number of people, and I would like to take the opportunity to thank them warmly.

First, I express my deepest gratitude to Professor Francesco Contino for proposing me to work on this inspiring topic and following my work since September.

Second, my gratitude further goes to Davide Tonelli, Researcher in the Institute of Mechanics, Materials and Civil Engineering, for his priceless guidance throughout the year. Without him, this work would not have come to an end.

I would also like to thank Professor Philippe Chevalier, for accepting to be a Reader of my Master Thesis.

Finally, words are not strong enough to express my warmest thanks to my family, for their continuous support day after day.

Chapter 2

Stochastic Model: Theoretical Review

The problem we want to solve has already a certain history in the literature. This chapter aims to look at key sources we have leveraged in this master thesis.

We start by a chronological review of important literature on the subject of hydrogen supply chain optimization. We then present two papers in more detail, that bring key concepts of stochastic programming to life in the context of hydrogen network optimization - and hence have largely inspired the development of our model. We then give a short overview of different approaches to uncertainty in the literature, not directly related to hydrogen supply chain planning. Finally, we introduce the theory behind multi-stage modeling, which forms the base of our own implementation.

2.1 Hydrogen Supply Chain (HSC)

In this first section, we present key papers in the hydrogen supply chain literature we have considered in our work - each offers a specific modeling or algorithmic approach. The references are presented chronologically.

(2003) Van den Heever and Grossmann [31]: This paper authors is often considered as pioneer in the field of HSC. Many would agree that the journey to optimization of hydrogen supply chain started with this work on short-term production planning. The model introduced in the paper produces a planning using demand forecasts with 12-hour time intervals. The model is deterministic, multi-period, mixed-integer, and introduces some non-linearity (MINLP). We see that a

very recurrent solution in the literature is to work around uncertainty and replace stochastic elements by 'good' estimates. This is not particular to HSC problems - numerous authors in supply chain management and planning approach uncertainty accordingly. Uncertainty parameters are approximated using forecasting strategies such as machine learning techniques and the values predicted are used directly as deterministic parameters. In some cases, this type of approach is a valid option. If the distribution of the parameters is highly concentrated, the solution will not change much. In most cases, however, this lack of precision requires post-process analysis to guarantee the consistency of the solution.

(2005) Hugo et al. [14]: The previous model was used for daily planning. Here, the authors address a long-range investment planning question - just like this master thesis. The model does not integrate any non-linearity (unlike Van den Heever and Grossmann); it is deterministic and multi-period (MILP). The objective of the authors is double: the model optimizes both the cost and the environmental impact of the solution. This comes up regularly in hydrogen infrastructure. Since the basic idea behind working with hydrogen is to reduce emissions, it makes sense to also integrate the environmental impact explicitly in objective function of the model. Note that our work focuses on prices and costs only.

(2006) Almansoori and Shah [2]: This paper uses a steady-state mathematical programming framework with deterministic demand for the minimization of the total cost. It's approach innovate by taking multiple part of the supply chain usually worked independently and try to optimize it as a whole.

(2008) Ingason et al. [16] and (2008) Lin et al. [21]: These two papers develop an MILP model to approach real cases study of hydrogen distribution planning in Iceland and Southern California, respectively. It use an integer representation to select the place of building new facilities.

(2008) Kim et al. [18] and Kim and Moon [19]: In these two papers, the authors use a stochastic approach to optimize a hydrogen network. They first propose a steady-state MILP model with demand uncertainty. The second paper extends the first results with a multi-objective optimization to consider both cost and safety.

(2009) Almansoori and Shah [4]: This paper is an extension of their 2006 work. The authors generalize their approach, from constant deterministic demand to multiple periods where each period has a different hydrogen demand pattern.

(2010) Guillén-Gosálbez et al. [12]: In this paper, the authors use a bi-level (master-slave) optimization approach where the master algorithm loop selects the technology that will be used while the slave algorithm then solves the simplified model using only the selected technology. The same strategy is used for our own approach with the Ant-Colony as meta-heuristic.

(2010) Sabio et al. (2010). [28]: In this important paper, the authors uses a bi-level meta-heuristic to handle integer variables via relaxation. They also introduce key concepts related to risk modelling. We present the paper and its approach in more detail in a dedicated section, hereafter.

(2012) Almansoori and Shah [3]: In this paper, the authors revisit their previous results by adding uncertainty to the demand. Clear source of inspiration of this thesis, the authors use a multi-stage model to treat uncertainty in hydrogen planning. His strategy of scenario-based with branching to capture uncertainty will form the base of our own model. Other papers took the same scenario-based stochastic programming approach to solve supply chain problems. For example, [5] introduces a MILP model in the context of hydrocarbon planning.

(2012) Agnolucci et al. [1]: In this paper, the author made an analysis of the optimization of hydrogen infrastructure literature and note in particular the necessity of a more rigorous treatment of uncertainty in demand.

(2019) Woo and Kim [32]: In this paper, finally, the authors use a genetic algorithm meta-heuristic to solve a MINLP formulation of the problem. They are not the only one to work on meta-heuristic as **Jesus et al. [27]** were also working on genetic algorithm at the same time. Our work follows a similar approach by introducing another type of evolutionary meta-heuristic into our algorithm, the Ant-Colony algorithm.

2.2 Two foundational papers in stochastic programming for HSC

In this section, we present two papers we largely leverage in this master thesis, and main lessons / implications for our work. Before that, we start with a short reminder of basic notations used later in the chapter.

Stochastic Programming: Basic Notations

- Ω : Sample-space, set of all possible scenarios
- 2^Ω : Set of all possible subsets of Ω
- ω : One possible resolution for the future, one possible scenario, with $\omega \in \Omega$; for each scenario correspond a specific value for each of the random parameters
- $\xi(\omega)$: Set of random parameters for scenario ω ; a sample is a particular value for the parameters and depends on the scenario considered
- Random variables can be described by their cumulative distribution function: $F_\xi(\epsilon) = P[\xi \leq \epsilon] = P[\omega | \xi(\omega) \leq \epsilon]$; in case of discrete random variable, we note this time $p_\xi(\xi_k) = P[\xi = \xi_k]$ with $\sum_k p_\xi(\xi_k) = 1$
- $f(x, \xi)$: Constraints and objectives can be expressed as function of those random variables and become random variables as well.

A Sabio et al. - Risk control of hydrogen supply

As explained earlier, when working with uncertainty, there is a certain observable preference in the literature to work with deterministic models as they are much simpler to solve. In order to do so, one can use the 'most-likely' approach and replace stochastic parameters by their expected values.

The objective of this paper is double but purely financial: when managing risk, the idea is to model a trade-off between financial risks and expected costs. Diverse methods exist. The one taken here is multi-objective. It minimize risks at the same time as costs in a combined objective function.

Working with uncertainty is indeed working with two concepts at the same time. First there is the robustness of the solution. The robustness can be any measure that defines the quality of a solution. It can typically be its performance when considering all possible cases for all realizations of the parameters. Secondly there is the efficiency of the solution which is the impact of the uncertainty on the objective function value. This is often modeled using an expectation of the costs or one of its derivatives. We have thus to work on both dimensions.

To compute the expected value, it is necessary to know the probability distribution of the random variable $\xi(\omega)$

$$\mathbb{E}_\xi[f(x, \xi)] = \int_\Omega f(x, \xi(\omega)) d\omega.$$

However, it often happens that this information is unknown. If some example of resolution for the parameters is known (e.g. $\xi = \{\xi_1, \dots, \xi_N\}$), the Monte-Carlo method can be used to compute an estimation of the expected value:

$$\mathbb{E}\xi[f(x, \xi)] = \sum_{i=1}^N f(x, \xi_i).$$

Risks can be modeled in different ways, as detailed here:

- **Chance constraint programming** Sometimes, probabilities can be used directly in models. Instead of requesting the constraints to hold, all the time, certain constraints have to be respected in α percent of the cases, where α is called 'confidence level'. If the joint probability distributions of the parameters are known (or good enough approximations), one can directly use the probability computation to build a new constraint:

$$P[g(x, \xi) \leq 0] = \int_{-\inf}^0 dG(\xi) \geq \alpha.$$

To check if such a condition is respected, a heuristic method can be used to generate possible values ξ_k of the random parameters:

$$P[g(x, \xi) \leq 0] \approx \frac{|\{g(x, \xi_k) \leq 0\}_{k=1}^N|}{N} \geq \alpha.$$

- **Dependent-chance programming (DCP)** This method also makes use of probabilities in the model to handle stochastic constraints, however this time using the probability in the objective function directly:

$$\max_x Pr[g_i(x, \xi) = c_i]_{i \in I}.$$

This is the choice of Sabio et al to define the risk. They want to maximize the probability to surpass a target. With f the objective function, d the design of the supply chain and ζ a predefined cost target, the risk can be written:

$$Risk(d, \zeta) = P[f(d) > \zeta].$$

This risk measure is then used with the expected cost to formulate a bi-objective model.

With some complex problems, there can be more than one of these probability measures to maximize. We will note $\{g_i(x, \xi) = c_i\}_{i \in I}$ the set of constraints

to maximize. Each constraint corresponds to a certain event. Unfortunately the probability of those events are linked together - they are not independent. In this method, the possible set of solutions is not convex and the number of constraints which can be respected at the same time is not known. The goal with the multi-objective model is to maximize the number of events respected concurrently:

$$\max_x |\{Pr[g_i(x, \xi) = c_i]\}_{i \in I}|.$$

Other methods can be used such as a weighted sum of objectives to give more or less importance to certain events:

$$\max_x \sum_{i \in I} \gamma_i Pr[g_i(x, \xi) = c_i].$$

Other extensions of this method exist such as 'dependent-chance goal programming' (DCGP) when the problem becomes really complex. In [23], a simulation genetic algorithm is proposed to solve the problem.

- **Min-max/Worst case scenario** Finally, a very simple yet very efficient strategy is to minimize the cost for the worst case scenario. This scenario can be defined by direct knowledge of our problem or using a min-max approach. When working with risk, we are sometimes only interested in worst case situations. We address this by a min-max analysis in the objective function. Using the dual problem, we can transform the maximization into a minimization and solve a single minimisation problem:

$$\max_{\xi} \min_x f(x, \xi) :$$

Note that working with worst case scenario could lead to an over-conservative solution with a lot of wasted resources in most of the scenarios.

Different sources of risks and other metrics can be integrated and optimized together through multi-objective formulations. In our case, for example, we will minimize the expected cost of the network design and the risk. The authors present two such formulations:

- **Weighted sum:** The simplest method is to sum up the different objectives. Predefined parameters λ_i give relative weights to the different objectives:

$$\begin{aligned} \min \quad & f(d) - \sum_{i \in I} \lambda_i g_i(d) \\ \text{s.t.} \quad & d \in \mathbb{D} \end{aligned}$$

with the different $g_i(d)$ being convex for the objective function to be convex, too.

- **ϵ -constraint:** For this method, the resolution is launched several times using only one of the objective while the second is set with confidence levels $\epsilon \in [\underline{\epsilon}, \bar{\epsilon}]$. The solution is of Pareto type. Contrary to the weighted sum, this method also works when $g(d)$ is non-convex:

$$\begin{aligned} \min \quad & E[f(d)] \\ \text{s.t.} \quad & d \in \mathbb{D} \\ & g(d) \leq \epsilon \end{aligned}$$

The treatment of uncertainty always brings a notable complexity increase in the model. The paper proposes a heuristic approach organised in two levels. First, a master algorithm solves a relaxation of the problem by considering integer as continuous variables. Second, a slave algorithm takes the solution from the master, rounds the relaxed integers and takes those values as parameters to solve the model once again. This approach is based on the observation that for some problems, the integers that represents the number of facilities to be established in a region takes particularly large values which means that the difference between relaxed and classic solutions is very high. It is however not really the case for decision variables, represented as Booleans, that define whether certain decisions are made or not. Note that this is exactly the type of variables we will have to formulate in our model.

In this paper, finally, uncertainty is introduced for the costs and is thus only present in the objective function but not in constraints and the demand is always exactly filled.

B Cantu et al. - Meta-heuristic bi-level optimization

One of the main sources of inspiration for this master thesis is the paper of Cantu et al [8]. The goal of the authors was to study hydrogen supply chain planning. In order to do so, they proposed to discretize the space into a grid, with production capacity and demand defined for each node of the grid. The model used was an MILP including a set of combinatorial variables to make the decisions of where to place production facilities and transport routes on the grid.

The complexity brought by the MILP formulation was tackled using a bi-level optimization algorithm. The master program handles the combinatorial part of the problem. The solution it gives represents possibilities to where and when to build new capacities. The slave algorithm then takes solutions from the master as parameters and solves the remaining linear problem to produce a feasible solution. This kind of heuristic to solve the problem in two rounds will also be part of our

own model.

The study presents a multi-objective strategy. The model tries at the same time to minimize daily cost of the network and environmental impact using a Pareto approach.

In the conclusion, the authors suggest different improvement paths. One of the suggestions is to introduce non-linearity in the problem formulation in order to get closer to real life situations. This is notably one of the objective of this master thesis.

2.3 Approaches to Uncertainty in the Literature

Uncertainty is a major factor when working on planning and management problems. While focusing on hydrogen supply chain, possible sources of uncertainty are numerous: equipment failure, natural disasters, market evolution, etc. We review different types of uncertainty here.

Endogenous vs exogenous: Uncertainty parameters are sometimes separated in two groups. First, there is endogenous uncertainty when the results are affected by the decision of the agent. Second, there is the exogenous uncertainty if the results are independent from the agent decisions. For example the model in [7] is a facility location with uncertain demand, but the demand is not considered completely exogenous as convenience of the facility impacts the demand. It works both on the cost and on the worst-case penalty function for unsatisfied demand. Our case could relate to this type of analysis as the future hydrogen demand could be linked to the investment made in the sector. We choose however to go for the simpler exogenous parameters.

Risk prevention: Uncertainty can also have a form of low probability but high impact events. In this type of analysis, a particular attention is given to extreme scenarios. Typically, in [7], the authors try to solve the problem of facility protection against disruption (natural disaster or intentional attacks). In our case, we are interested in long-term planning and evolution, more than preparing for an eventual crash of the hydrogen market.

Tight constraint: The impact of uncertainty is not the same if it applies to the objective function or constraints. If we want the constraints to always hold, it would imply a certain computational cost to check all possible combinations of resolution for the uncertain parameters. As the problem is more constrained, it implies a loss

of quality since the space of available solutions is reduced. The solution will always be less good than if each scenario is considered as a deterministic problem.

Extreme scenarios can have a huge impact on the solution, and that even if they are very unlikely to happen. Typically, it occurs when working with strict robustness. If we want constraints to be always respected, extreme values will lead to very hard constraints. For example, if a demand could be theoretically infinite. In certain cases, extreme scenarios will be overlooked. Of course this method is not valid for situations where those scenarios, yet unlikely, could have devastating consequences as in the nuclear sector for instance.

For example, in [6], the authors work on the planning of water supply with unknown future conditions. The model developed has multiple objectives: quality and robustness. The supply condition is relaxed: supply has to be superior or equal to the demand in all scenarios, instead of requiring being tight.

2.4 Multi-Stage Modeling

When stochastic programming is the choice, two options come out. The first one is stochastic programming with recourse, while the other is robust optimization. We have already discussed ways to work with risk and robustness in the previous sections. Here we introduce recourse and multi-stage modeling (see e.g. [22] and [29]).

We have also seen that certain papers use multi-period models. In these approaches, a solution gives for each period the optimal set of decisions to make. A multi-stage is more adaptive since certain decisions depend on the results of the uncertainty parameters. The chosen solution could diverge after some rounds depending on the scenario we are in, and the initial decision of time 1 will take this into account. On the contrary, a multi-period model gives only one solution by time-step.

A Two-Stage Model and Recourse Actions

The simplest model of multi-stage programming is the 2-stage one, structured as follows:

- 'Here and now' variables: A first set of decisions $x \in \mathbb{R}^{n_1}$ is made before realization of the random parameters
- Then, uncertainty is unveiled and the true value of the $\xi(\omega)$ parameters gets known; from this point on, the scenario $\omega \in \Omega$ we are in is known

- 'Wait and see' variables: A second set of decision $y(\xi(\omega)) \in \mathbb{R}^{n_2}$ named recourse decision can be made, with perfect information; those are the 'recourse' actions.

A classic 2-stage linear problem with recourse can be written in two parts: first the minimization of the cost at the first stage and the expected cost for the next stages:

$$\begin{aligned} \min \quad & c^T x + \mathbb{E}[Q(x, y, \xi(\omega))] \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0, \end{aligned}$$

Followed by the expression of the cost in second stage as another minimization problem:

$$\begin{aligned} Q(x, \xi(\omega)) = \min \quad & q(\xi(\omega))^T y(\xi(\omega)) \\ & T(\xi(\omega))x + W(\xi(\omega))y(\xi(\omega)) = h(\xi(\omega)) \\ & y \geq 0 \end{aligned}$$

The constraints defined by the parameters $A \in \mathbb{R}^{m_1 \times n_1}$ and $b \in \mathbb{R}^{m_1}$ are deterministic as well as the c in the objective function. A part of the objective function is still uncertain and quantified by the parameters $q(\omega) \in \mathbb{R}^{n_2}$. The constraints defined by $T(\omega) \in m_2 \times n_1$ and $W(\omega) \in m_2 \times n_2$ depend on the state of the model. To assure that these constraints are respected, the recourse decision variables $y(\omega)$ will be necessary.

$Q(x, \xi(\omega))$ is the recourse function, the optimal value of the second-stage problem. When $\xi(\omega)$ has finite support, it means that there is a finite number of possible scenarios $\xi_1, \dots, \xi_K$, with respective probabilities $p_k \in (0, 1)$. The recourse function can be written $Q(x, y) = \mathbb{E}[Q(x, y, \xi(\omega))] = \sum Q(x, \xi_k)$. When the recourse matrix $W(\omega)$ is independent from ω , we say that the problem is 'with fixed recourse'. This property allows the use of certain specific algorithm that ease the resolution.

One can stack several two-stage structures for several successive time periods, thereby creating a multi-stage model. After each time-step, new information becomes available about outcoming results for this time-step while new decisions need to be made. These new decisions are themselves based on uncertainty parameters for the time-steps to come.

B Scenario Tree

The problem takes place in a finite sequence of time periods $t \in \{1, \dots, H\}$. The information is progressively unveiled one time period after the other. We define one node of the tree as state s . Each state s correspond both to: a time step t , a particular scenario for the uncertainty parameters, and a set of possible action A_t . These sets are typically vectors of random parameters and variables.

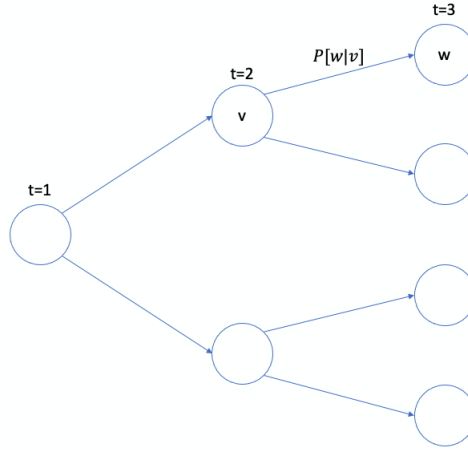


Figure 2.1: Tree of states for multi-stage formulation

The root of the tree corresponds to time $t = 1$ and state $s = 0$. To each state s corresponds one time period $t(s)$. We note $\mathcal{S}(s)$ the successors of state s ; they are the possible states following s in time $t(s) + 1$. The ancestors $\mathcal{C}(s)$ of state s are the states that form in the tree the path from the root to s . To each jump from a state to another corresponds a probability transition. It is the probability for a certain realization in the future knowing the current state of the model.

C Multi-Stage Formulation

The multi-stage is the adaption of the two-stage model, with several successive time-steps. Instead of one static solution, the model produces an adaptive solution, the which adaption cycle of which is defined by the time-steps. The stochastic formulation of the model goes as follows:

$$\begin{aligned}
\min \quad & c_1^T x_1 + \mathbb{E}[c_2(\omega)^T x_2(\omega) + \dots + c_H(\omega)^T x_H(\omega)] \\
\text{s.t.} \quad & A_1 x_1 = b_1 \\
& T_1(\omega)x_1 + W_2(\omega)x_2(\omega) = h_2(\omega) \\
& \vdots \\
& T_{t-1}(\omega)x_{t-1}(\omega) + W_t(\omega)x_t(\omega) = h_t(\omega) \\
& \vdots \\
& T_{H-1}(\omega)x_{H-1}(\omega) + W_H(\omega)x_H(\omega) = h_H(\omega) \\
& x_1 \geq 0, x_t(\omega) \geq 0 \quad t = 2, \dots, H.
\end{aligned}$$

D Quality metrics: EVP and VSS

As we have seen, by taking the uncertainty into account we made our model quite complex. Here we will see an usual way to test the usefulness of the approach.

We call Expected Value (EV) the solution of the minimization problem when the uncertain parameters ξ take their expected value $\bar{\xi}$. It is as we didn't took uncertainty into account, and just solve the system replacing uncertain parameters by their expectation. We can then take $\bar{x}(\bar{\xi})$ as optimal values of decision variables for this case.

$$EV = \min_x f(x, \xi) \rightarrow \bar{x}(\bar{\xi})$$

The Expected Value Problem (EVP) is the expectation of the objective function on all the possible scenario of ξ , but taking as decision variable the solution from the expected case $\bar{x}(\bar{\xi})$ all case.

$$EVP = \mathbb{E}_\xi[f(\bar{x}(\bar{\xi}), \xi)]$$

The Value of the Stochastic Solution (VSS) is then defined as the difference between the Expected Value Problem and the classic solution taking the uncertainty parameters into account:

$$VSS = \mathbb{E}_\xi[f(\bar{x}(\bar{\xi}), \xi)] - \min_x E_\xi[f(x, \xi)].$$

This usually serves to assess the quality of a solution. By comparison to the case when we ignore uncertainty, we can see how much we gain by playing on it. However it is hardly applicable in our situation. We will see in the model presentation that some of our constraints have to be tight. Those are also the main interest of working with uncertainty parameter and not only the final objective function.

Chapter 3

Model

Based on the literature review, we have chosen to model our problem using a combination of mixed-integer non-linear programming together with a multi-stage approach to integrate sources of uncertainty. In this chapter, we present our model and modeling options in detailed mathematical terms. We start with the presentation of the objective function, follow with an explanation of constraints taken into account and finally introduce detailed declarations of sets (search spaces), variables and parameters used.

Given the top-down presentation approach we have chosen in this chapter, we invite the reader to consider the latter part as reference to fully understand the first sections.

3.1 Key Modeling Choices

The model we are developing here is characterized by the presence of certain simplifications compared to a real case.

The hydrogen is usually produced, transported, stored in diverse forms, at different pressure levels. Those points are ignored in the current model, but our generic formulation allows the introduction of these new features. Moreover we do not consider storage facility planning in the current problem.

The model works using energy carriers as generic variables. The solution of the model gives an optimal design for a chemical supply chain which includes a predefined set of carriers (e.g. hydrogen, carbon dioxide, natural gas) and the electric power system. Our model also gives the coupling between the different

energies using production capacities as transformation ways between energy carriers.

The solution describes the location, sizing and operation of production and transport technologies that minimize the cost and maximize the reliability of the system. The optimal solution is space- and time-resolved and is determined for different uncertain scenarios.

This requires knowledge of the consumption of the carriers of interest (e.g. energy demands), of resource (e.g. renewable energy sources) and technology availability, and of technology performance parameters. These different points are covered in the latter sections of this chapter.

3.2 Objective Functions

The objective function to optimize is the expected cost of the system for all possible states:

$$J = E_s[J_s^{imp} + J_s^{exp} + J_s^{cap} + J_s^{op}]$$

It is calculated as the sum for all states, of the cost in each state multiplied by the probability to find the system in that state:

$$J = \sum_{s \in \mathcal{S}} p_s (J_s^{imp} + J_s^{exp} + J_s^{cap} + J_s^{op}).$$

For each state, the total cost is the sum of 4 terms:

- **Carrier import:** cost of importing all the necessary carriers of energy to fuel the system

$$J_s^{imp} = \sum_{i \in \mathcal{N}} \sum_{c \in \mathcal{C}} U_{cis} u_{cis}$$

- **Carrier export:** gain from exporting all the surplus of carriers

$$J_s^{exp} = \sum_{i \in \mathcal{N}} \sum_{c \in \mathcal{C}} V_{cis} v_{cis}$$

- **CAPEX:** cost of installing new production and transport capacities ($dS_{q,i,s}$ and $dC_{q,i',s}$). The cost is non-linear and depends on the already installed

capacities ($S_{q,i,s}$ and $C_{q,i,i',s}$). For the transport, it also depends on the distances between the nodes we want to link ($l_{q,i,i'}$)

$$J_s^{cap} = \sum_{i \in \mathcal{N}} \sum_{q \in \mathcal{Q}_p} c_{q,s}^p(S_{q,i,s}, dS_{q,i,s}) + \sum_{s \in \mathcal{S}} \sum_{i, i' \in \mathcal{N} \times \mathcal{N}} \sum_{q \in \mathcal{Q}_t} c_{q,s}^t(C_{q,i,i',s}, dC_{q,i,i',s}, l_{q,i,i'})$$

- **OPEX:** cost to produce the output flows $P_{c,q,i,s}$ of carriers, and the cost of operating the flows $F_{c,q,i,i',s}$ through the transport network. For the transport, this also depends on distances between the nodes we want to link

$$J_s^{op} = \sum_{i \in \mathcal{N}} \sum_{q \in \mathcal{Q}_p} o_{q,s}^p(P_{c,q,i,s}) + \sum_{s \in \mathcal{S}} \sum_{i, i' \in \mathcal{N} \times \mathcal{N}} \sum_{q \in \mathcal{Q}_t} o_{q,s}^t(F_{c,q,i,i',s}, l_{q,i,i'}).$$

3.3 Constraints

In this section we describe the constraints of the optimization problem, i.e. mass balance, conversion and transport technology performance, and carrier availability.

A Mass Balance

At all states $s \in \mathcal{S}$, at every node $i \in \mathcal{N}$, for all carriers $c \in \mathcal{C}$, the corresponding energy balance must be satisfied in all nodes $i \in \mathcal{G}$:

$$\begin{aligned} & \sum_{q \in \mathcal{Q}_p} (P_{c,q,i,s} - G_{c,q,i,s}) \\ & + \sum_{q \in \mathcal{Q}_t} \sum_{i' \in \mathcal{N}, i' \neq i} (F_{c,q,i',i,s} - f_q(F_{c,q,i',i,s}, l_{q,i,i'}) - F_{c,q,i,i',s}) \\ & + U_{c,i,s} - V_{c,i,s} - d_{c,i,s} \\ & = 0. \end{aligned} \tag{3.1}$$

In the previous equation:

- $F_{c,q,i',i,s} - f_q(F_{c,q,i',i,s}, l_{q,i,i'}) - F_{c,q,i,i',s}$ represent respectively the import (first 'F') and export (second 'F') of carrier c at node i and state s as well as loss of carrier c ('f') during the transport. It is calculated using a sum over all edges i, i' connected to the node, and all transport technologies $q \in \mathcal{Q}_t$ used by the system
- The difference $P_{c,q,i,s} - G_{c,q,i,s}$ represents the output of carrier c at node n from local production or input to local production
- $U_{c,i,s}$ and $V_{c,i,s}$ respectively model the import and export of carrier c for the node i at state s
- $d_{c,i,s}$ represents the demand of carrier c at node i at state s that we are trying to fill.

B Conversion Technology Performance

Limited technology availability: The size of conversion technology q is bounded by the maximum potential installation allowed in node i , $s_{qi,t}^{\max}$. For all conversion technologies $q \in \mathcal{Q}_p$, for all nodes $i \in \mathcal{N}$, for all state $s \in \mathcal{S}$, this is expressed as:

$$S_{qis} \leq s_{qit_s}^{\max} \quad (3.2)$$

with t_s being the time step corresponding to state s .

Each conversion technology possesses a reference carrier, such that 1 capacity unit of this technology corresponds to 1 unit of this carrier (see eq. (3.4)). The inputs and outputs of other carriers are calculated relatively to the reference one.

Input/output flows: The performance of a conversion technology q , using carriers $c \in \mathcal{J}_h$ as input, and producing carriers $c' \in \mathcal{J}_h$ as output, is modeled through the function $\eta_{q,c,c'}$, which quantifies the efficiency of technology q in converting carrier c into carrier c' . This is expressed as:

$$P_{c',q,i,s} = \eta_{q,c,c'}(S_{q,i,s}, G_{c,q,i,s}, P_{c',q,i,s})G_{c,q,i,s} \quad (3.3)$$

with the reference output being a non-negative quantity constrained by the installed size $S_{q,i,s}$ of conversion technology q in node i , at state s and the minimum load of the conversion technology.

The equation links the input flow $G_{c,q,i,s}$ of carrier c and the output flow $P_{c',q,i,s}$ of each node i in state s for a particular technology q transforming carrier c into carrier c' .

Loading: If technology q is in operation ($y_{q,i,s} = 1$) at state s and node i , the flow of the reference carrier ($s = \mathcal{R}(q)$) is subject to a lower bound value different from zero, and an upper bound determined by the installed capacity. Otherwise it is constrained to zero independently of the installed capacity, when the technology is not operating ($y_{q,i,s} = 0$):

$$p_h^{\min} y_{q,i,s} S_{q,i,s} \leq P_{c,q,i,s} \leq S_{q,i,s} y_{q,i,s}. \quad (3.4)$$

The flows for the other carriers are determined using equation 3.3 and suffer from the same limitation, up to the $\eta_{q,c,c'}(\cdot)$ parameter.

Capacity expansions: The capacity of the conversion technology q at node i in state s can be defined as the sum of the capacity expansions over a number of time-steps equal to the technology lifetime t_q^l . Since the installation of capacity is state-dependent we have to take the right predecessor states to compute this sum. By referring to $dS_{q,i,s'}$ as the capacity expansion at state s' , the capacity $S_{q,i,s}$ is defined as:

$$S_{q,i,s} = \sum_{s' \in \mathcal{C}(s), t_{s'} \in \{t_s - t_q^l, t_s\}}^t dS_{q,i,s'}. \quad (3.5)$$

Furthermore, no deconstruction is allowed over the lifetime of the conversion technology:

$$dS_{q,i,s} \geq 0. \quad (3.6)$$

Min/max capacity expansion: The increase in size has to be at least equal to δs_q^{min} and at most equal to δs_q^{max} :

$$\delta s_q^{min} z_{q,i,s} \leq dS_{q,i,s} \leq \delta s_q^{max} z_{q,i,s} \quad (3.7)$$

where the binary decision variable $z_{q,i,s}$ describes whether the capacity of the conversion technology q is expanded at the state s ($z_{q,i,s} = 1$), or not ($z_{q,i,s} = 0$).

A capacity expansion can happen multiple times during the lifetime of a plant, but requires an amount of time steps t_q^e for conversion technology q . This condition is enforced with the following inequality:

$$\sum_{s' \in \mathcal{C}(s), t_{s'} \in \{t_s - t_q^e, t_s\}}^t z_{q,i,s'} \leq 1 \quad (3.8)$$

which imposes that only one expansion can happen during the expansion time t_q^e , for the series of states that precede state s .

C Transport Technology Performance

Limited technology availability: The size of transport technology $q \in \mathcal{Q}_t$ is bounded by the maximum potential installation $s_{q,i,t}^{max}$ at location i at time t . For all transport technologies $q \in \mathcal{Q}_t$, for all connections $i, i' \in \mathcal{N}$, for all state $s \in \mathcal{S}$, this is expressed as:

$$C_{q,i,i',s} \leq s_{q,i,t_s}^{max} \quad (3.9)$$

with t_s being the time step corresponding to state s .

Input/output flows: The flow $F_{c,q,i,i',s}$ of carrier c through transport technology q between nodes i and i' at state s is constrained by the installed size $C_{q,i,i',s}$ and its minimum operation load $p_k^{\min}$. If the technology is in operation ($y_{q,i,i',s} = 1$) at state s , the flow is subject to a lower bound value different from zero, otherwise is constrained to zero independently of the installed capacity, when the technology is not operating ($y_{q,i,i',s} = 0$):

$$p_k^{\min} y_{q,i,i',s} C_{q,i,i',s} \leq F_{c,q,i,i',s} \leq C_{q,i,i',s} y_{q,i,i',s}. \quad (3.10)$$

Capacity expansion: Similarly to the conversion technologies, the installed capacity of transport technology q is expressed as sum of the capacity expansions over all preceding states $\mathcal{C}(s)$, that were in the right time period relatively to the technology lifetime t_q^l of the considered technology:

$$C_{q,i,i',s} = \sum_{s' \in \mathcal{C}(s), t_{s'} \in \{t_s - t_k^l, t_s\}} dS_{k,i,i',t',s} \quad (3.11)$$

where we sum the capacity expansions.

Here too, no deconstruction is allowed over the lifetime of the transport technology; therefore,

$$dC_{q,i,i',s} \geq 0. \quad (3.12)$$

Min/max capacity expansion: The increase in size has to be at least equal to $\delta s_q^{\min}$ and at most equal to $\delta s_q^{\max}$:

$$\delta s_k^{\min} z_{q,i,i',s} \leq dC_{q,i,i',s} \leq \delta s_q^{\max} z_{q,i,i',s} \quad (3.13)$$

where the binary decision variable $z_{q,i,i',s}$ describes whether the capacity of the transport technology q is expanded at the state s ($z_{q,i,i',s} = 1$), or not ($z_{q,i,i',s} = 0$).

A capacity expansion can happen multiple times during the lifetime of a technology, but requires an amount of time steps t_q^e for transport technology q . This condition is enforced with the following inequality:

$$\sum_{s' \in \mathcal{C}(s), t_{s'} \in \{t_s - t_k^e, t_s\}} z_{q,i,i',s'} \leq 1 \quad (3.14)$$

which imposes that only one expansion can happen during the expansion time t_q^e .

D Carrier Availability

Depending on node, state, and scenario, the carrier availability can be limited. This limitation is not described by an uncertain parameters and is thus time-dependent. Therefore, for each node, state, the import $U_{j,i,s}$ of carrier j in node i , used in any technology in the node cannot exceed its maximum availability a_{j,i,t_s} , with $t_s \in \mathcal{T}$ being the time corresponding to state s in the state tree. The limited availability of carrier $j \in \bar{\mathcal{L}}$ in node $i \in \mathcal{N}$ at state s is expressed as follows:

$$U_{j,i,s} \leq a_{j,i,t}. \quad (3.15)$$

3.4 Formal Declarations

In this section we detail out all declarations needed for the model: states, sets, input parameters, input functions, decision parameters.

A State

A state defines both a time-step and a particular possible scenario for the evolution of the system and parameters. The choices made at a state s will have an impact in the time-steps that follow on all its successors $\mathcal{S}(s)$. Almost all states also possess a predecessor state, which defines the system in the previous time period but also the decisions made there as part of the solution. We can go back in time by successively taking the predecessor of each state until we arrive at state zero. It is the only state that has no predecessor. This will draw a path in the state tree with all the states that the system have been through to arrive at state s . We define this set as $\mathcal{C}(s)$.

Figure (3.1) illustrates part of the tree of states. State 4 as two predecessors, namely $\mathcal{C}(4)$, i.e. states 0 and 1. It also has two successors, namely $\mathcal{S}(4)$, i.e. states 9 and 10.

B Sets

The two tables below define all value sets necessary to formulate our model mathematically.

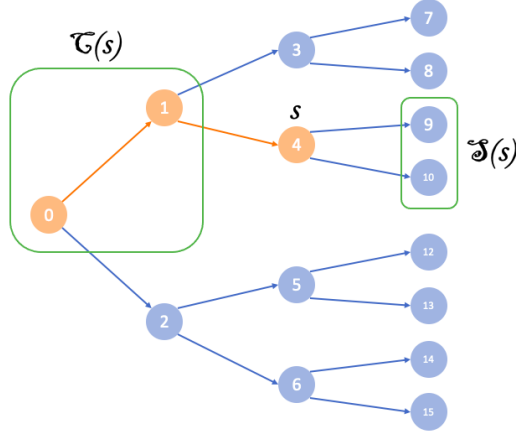


Figure 3.1: Tree of states with the set of successors $\mathcal{S}(4)$ and the set of predecessors $\mathcal{C}(4)$ for a particular state 4

Table 3.1: Main sets

Name	Description	Index notation
$\mathcal{N}$	nodes (<i>setNodes</i>)	i
$\mathcal{N}_c \subset \mathcal{N}$	nodes that represent regions in the country	i
$\mathcal{N}_n \subset \mathcal{N}$	nodes that represent neighbouring countries	i
$\mathcal{N}_p \subset \mathcal{N}$	nodes that represent the possible import through one port on the North Sea (only one node)	i
$\mathcal{C}$	energy carriers (<i>setCarriers</i>)	c
$\mathcal{T}$	time periods (<i>setTimeSteps</i>)	t
$\mathcal{S}$	state in the stochastic scenario tree (<i>setStates</i>)	s
$\mathcal{Q}$	technologies	q
$\mathcal{Q}_p \subset \mathcal{Q}$	hydrogen production technologies (<i>setConversionTechnologies</i>)	q
$\mathcal{Q}_t \subset \mathcal{Q}$	hydrogen transport technologies (<i>setTransportTechnologies</i>)	q

Table 3.2: Auxiliary sets - Note: Parentheses indicate a dependency of elements of the set on elements of another set

Sets	Description
$\mathcal{D}((\mathcal{Q}_p \mathcal{Q}_t), \mathcal{N})$	combination of technologies with nodes for conversion technologies and pairs of nodes (edges) for transport technologies (<i>setTechnologyLocation</i>)
$\overline{\mathcal{F}}(\mathcal{Q}_p, \overline{\mathcal{J}}(\mathcal{Q}_p))$	combination of technologies with the respective input carriers (<i>setInputCarriersTechs</i>)
$\underline{\mathcal{F}}(\mathcal{Q}_p, \underline{\mathcal{J}}(\mathcal{Q}_p))$	combination of technologies with the respective output carriers (<i>setOutputCarriersTechs</i>)
$\mathcal{G}(\mathcal{Q}_t, \mathcal{Q}(\mathcal{Q}_t))$	combination of technologies with the respective carriers (<i>setTransportCarriersTech</i>)
$\overline{\mathcal{J}}(\mathcal{Q}_p)$	input carriers of conversion and transport technologies (<i>setInputCarriers</i>)
$\underline{\mathcal{J}}(\mathcal{Q}_p)$	output carriers of conversion and transport technologies (<i>setOutputCarriers</i>)
$\mathcal{M}(\mathcal{Q}_t)$	transport carriers of transport technologies (<i>setTransportCarriers</i>)
$\mathcal{P}(\mathcal{Q}_p)$	dependent carriers of conversion and transport technology (<i>setDependentCarriers</i>)
$\mathcal{R}(\mathcal{Q}_p \mathcal{Q}_t)$	reference carriers of conversion and transport technology (<i>setReferenceCarriers</i>)
$\mathcal{S}(\mathcal{S})$	combination for each state and of all possible successor state in time t+1 (for a state of time t) (<i>setSuccessorStates</i>)
$\mathcal{C}(\mathcal{S})$	combination for each state of all its predecessors (all state from time 1 to the state in question, see figure 3.1) (<i>setPredecessorStates</i>)

C Input Parameters

Our model includes a number of input parameters. Among those, 3 inputs have a stochastic nature, are defined as probability distributions and models as sets of scenarios - we note the space of possible scenarios Ω . The sources of uncertainty in the model are summarized below.

1. $d_{cis}(\omega)$: Demand for carriers
2. $u_{cis}(\omega)$: Import costs

3. $v_{cis}(\omega)$: Export costs.

The table below formalizes all input parameters.

Table 3.3: Input parameters

Name	Description	Domain	U. m.	Stoch.
d_{cis}	demand of carrier c , at node i , in the model state s (<i>demandCarrier</i>).	$\mathbb{R}^+$	MWh/yr	YES
u_{cis}	import price of carrier c , in node i , at state s (<i>importPriceCarrier</i>).	$\mathbb{R}^+$	k€/MWh	YES
v_{cis}	export price of carrier c , in node i , at state s (<i>exportPriceCarrier</i>).	$\mathbb{R}^+$	k€/MWh	YES
a_{cit}	availability of carrier c , at node i , at time t (<i>availabilityCarrier</i>).	$\mathbb{R}^+$	MWh/yr	NO
p_q^{min}	minimum load of conversion/transport technology q (<i>minLoadConversionTechnology</i>).	$\mathbb{R}$ & $[0, 1]$		NO
s_{qit}^{max}	maximum capacity of conversion/transport technology h in node i (<i>maxPotentialCapacity</i>).	$\mathbb{R}^+$	MW	NO
δs_q^{min}	minimum capacity allowed capacity expansion of conversion/transport technology q		MW	NO
δs_q^{min}	maximum capacity allowed capacity expansion of conversion/transport technology q		MW	NO
$l_{qii'}$	Distance between node i and i' for transport technology q (<i>distance</i>).	$\mathbb{R}^+$	km	NO
$i_{qii'}$	halved cost per unit distance of the nondirectional capacity connecting nodes i and i' for transport technology q (<i>costPerDistance</i>).	$\mathbb{R}^+$	EUR/km	NO

$m_{qii'}$	halved maintenance cost of per unit distance of the nondirectional capacity connecting nodes i and i' for transport technology q (<i>maintenanceCostPerDistance</i>)	$\mathbb{R}^+$	EUR/(km yr)	NO
k_q	loss coefficient per unit of distance for transport technology q (<i>lossFlow</i>)	$\mathbb{R}^+$	1/km	NO
t_q^l	lifetime of conversion/transport technology q (<i>lifetimeConversionTechnology</i> , <i>lifetimeTransportTechnology</i>)	$\mathbb{N}$	yr	NO
t_q^e	expansion time required for conversion/transport technology (<i>expansionConversionTechnology</i> , <i>expansionTransportTechnology</i>) q	$\mathbb{N}$	yr	NO
μ_{qi}	Maintenance cost of a technology q at node i .	$\mathbb{R}^+$	k€/MW	NO
t_s	Time corresponding to the state s	$[0, H]$	yr	NO
p_s	Probability corresponding to the state s	$]0, 1]$		NO

D Input (Non-Linear) Functions

Table 3.4: Input (non-linear) functions

Function	Input	Description	Codomain	U.m.
c_q^p	$S_{q,i,s}, dS_{q,i,s}$	function mapping the size of the conversion technology q and the planned new capacity to its installation cost (<i>capex</i>)	$\mathbb{R}_+$	EUR
o_q^p	$P_{c,q,i,s}$	function mapping the output stream of the conversion technology q to its operational cost under scenario s	$\mathbb{R}_+$	EUR

c_q^t	$C_{q,i,i',s},$ $dC_{q,i,i',s},$ $l_{q,i,i'}$	function mapping the size and the distance between nodes of the transport technology q to its installation cost (<i>capex</i>)	$\mathbb{R}_+$	EUR
o_k^t	$F_{j,q,i,i',s},$ $l_{q,i,i'}$	function mapping the flow and distance of the transport technology q to its operational cost	$\mathbb{R}_+$	EUR
$\eta_{q,c,c'}$	$S_{q,i,s},$ $G_{c,q,i,s},$ $P_{j,q,i,s}$	function mapping the size, input and output stream of conversion technology q converting carrier c to carrier c'	$\mathbb{R} \cap [0, 1]$	-
f_q	$F_{c,q,i,i',s},$ $l_{q,i,i'}$	function mapping the flow and transport distance of carrier c with transport technology q between node i and i' to the transport losses (<i>carrierLoss</i>)	$\mathbb{R} \cap [0, 1]$	-

E Decision Parameters

Since the resolution of the parameters will imply different decisions as the system goes through each time-step, all decision variables are state-dependent.

Table 3.5: Decision parameters

Variable	Description	Domain	U. m.
$y_{q,i,s}$	scheduling of operation of conversion technology q at node i and state s (<i>selectConversionTechnology</i>)	$\{0,1\}$	-
$z_{q,i,s}$	expanding the capacity of conversion technology q at node i and state s , subject to installation time t_q^i (<i>expandConversionTechnology</i>)	$\{0,1\}$	-
$y_{q,i,i',s}$	scheduling of operation for transport technology q from node i to node i' at state s (<i>selectTransportTechnology</i>)	$\{0,1\}$	-
$z_{q,i,i',s}$	expanding the capacity of transport technology q from node i to node i' at state s (<i>expandTransportTechnology</i>)	$\{0,1\}$	-
$S_{q,i,s}$	capacity of conversion technology q at node i and state q (<i>capacity</i>)	$\mathbb{R}_+$	MW

$C_{q,i,i',s}$	capacity of transport technology q from node i to node i' at state s (<i>capacity</i>)	$\mathbb{R}_+$	MW
$dS_{q,i,s}$	capacity increase of conversion technology q at node i and state s (<i>builtCapacity</i>)	$\mathbb{R}_+$	MW
$dC_{q,i,i',s}$	capacity increase of transport technology q from node i to node i' at state s (<i>builtCapacity</i>)	$\mathbb{R}_+$	MW
$G_{c,q,i,s}$	input stream of carrier c into conversion technology q at node i and state s (<i>inputFlow</i>)	$\mathbb{R}_+$	$\frac{\text{MW h}}{\text{yr}}$
$P_{c,q,i,s}$	output stream of carrier c from conversion technology q at node i and state s (<i>outputFlow</i>)	$\mathbb{R}_+$	$\frac{\text{MW h}}{\text{yr}}$
$F_{c,q,i,i',s}$	flow of carrier c through transport technology q from node i to node i' and state s (<i>carrierFlow</i>)	$\mathbb{R}_+$	$\frac{\text{MW h}}{\text{yr}}$
$U_{c,i,s}$	import of carrier c into node i from the grid at state s (<i>importCarrierFlow</i>)	$\mathbb{R}_+$	$\frac{\text{MW h}}{\text{yr}}$
$V_{c,i,s}$	export of carrier c from node i to the grid at state s (<i>exportCarrierFlow</i>)	$\mathbb{R}_+$	$\frac{\text{MW h}}{\text{yr}}$

Chapter 4

Belgian Case Study: Key Dimensions and Design Choices

As introduced in the first chapter, this master thesis aims to apply stochastic modeling to define optimal future hydrogen infrastructures at country or regional level, and to test our developments on a case study, i.e. Belgium.

We propose to use a multi-stage, multi-scenario approach which allows to tackle both the evolution of our network over time but also the different sources of uncertainty in our model, namely the demand for carriers and their import prices.

The model defined in the previous chapter was kept generic on purpose. In the next sections, we present the Belgian case study, key sources used to define model parameters as well as our main design choices to implement the model.

4.1 Discrete Time Evolution

The time-steps of our multi-stage model are relatively large, i.e. around a decade ($H = \{2022, 2030, 2040, 2050\}$). We consider that new capacities are built during the time comprised between two time-steps and become operational just at the time-step directly following the one where the decision was made. Clearly, shorter time-steps can easily be introduced in a second stage, to generate more precise solutions (esp. from an industrial and construction point of view).

4.2 Topography as a Network

The country, Belgium, and its neighbors are represented on figure 4.1. To keep a reasonable level of complexity, we propose to consider the geocode standards NUTS-2 and NUTS-0 from the Nomenclature of Territorial Units for Statistics of the European Union.

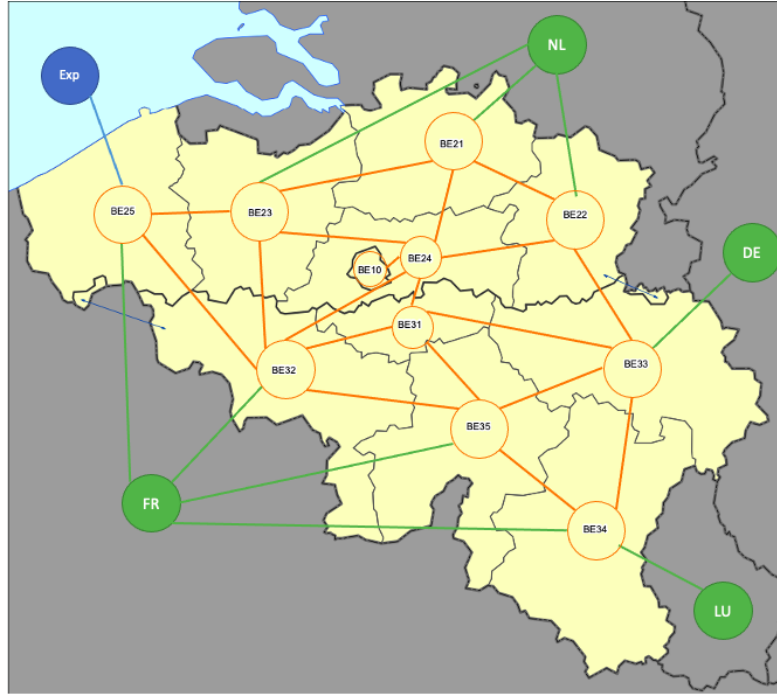


Figure 4.1: NUTS-2 Belgium

- Within Belgium, we use the province as the smallest geographic unit (we could obviously consider large agglomerations or even municipalities in a second stage, with the same model and implementation design): Brussels (BE10), Antwerp (BE21), Limburg (BE22), East Flanders (BE23), Flemish Brabant (BE24), West Flanders (BE25), Walloon Brabant (BE31), Hainaut (BE32), Liège (BE33), Luxembourg (Belgium) (BE34), Namur (BE35).
- Each of our neighboring countries is modeled as a single node aggregating all the data (volumes and costs): France (FR), Germany (DE), the Netherlands (NL)
- Finally, a fictitious point represents the connection to the rest of the world by the North Sea (NS) and Belgian port capabilities.

The final network takes the form as presented in figure ?? . Each region has a topographic position defined by its center. Distances used to compute the transport costs between regions are euclidean distances between pairs of nodes.

There is in total 16 nodes in this network which means a potential of 120 edges and lots of combinations to explore. To simplify the model and, later, reduce

execution time to find optimal solutions, we propose to reduce the number of edges. Indeed, directly connecting two distant regions is not practically feasible and hence 'long-stances edges' will be eliminated upfront from the set of solutions (otherwise, they would always be preferred). Practically available edges will be directly specified as parameters for the solution. With this, we keep the network-based approach for solutions while removing numerous variables and reducing the resulting complexity.

4.3 Hydrogen Demand

Regarding carrier demand, we are mainly interested in hydrogen. For this reason, the demand for other carriers will be set to 0 at all time. Once again, other carriers could be treated by the model in a second stage to generate more realistic solutions, at the expense of execution time to generate optimal solutions.

Our model considers an uncertain demand at each node. We will use an approximation of this demand provided in the Hydrogen Roadmap Europe [15]. Based on inputs from "17 member companies active in hydrogen and fuel cell technologies", this source provides for each region an estimation of the demand for different scenarios. The most ambitious scenario corresponds to EU achieving the 2-degree stabilisation target. In this situation, hydrogen represents 24% of the total energy demand in Europe in 2050 (approx. 2,250 TWh). This demand estimation includes applications for power generation and buffering, transportation, heating and power for buildings, and industrial energy and feedstock needs.

Following the same idea, the demand will be addressed as percentage of total demand. The demand is modeled as uniform distribution from the worst to the best solution. We keep the idea of a global dynamics between successive states. If a node has a lower (resp. higher) value than another node, the children of the first node will also have a lower (resp. higher) value than the children of the other node - see figure 4.2.

As for today, hydrogen represent only 2% of the total energy demand. For the sake of simplification we will consider that we have to start from scratch: in the initial state, the network does not have any production/transport capacity and demand is null.

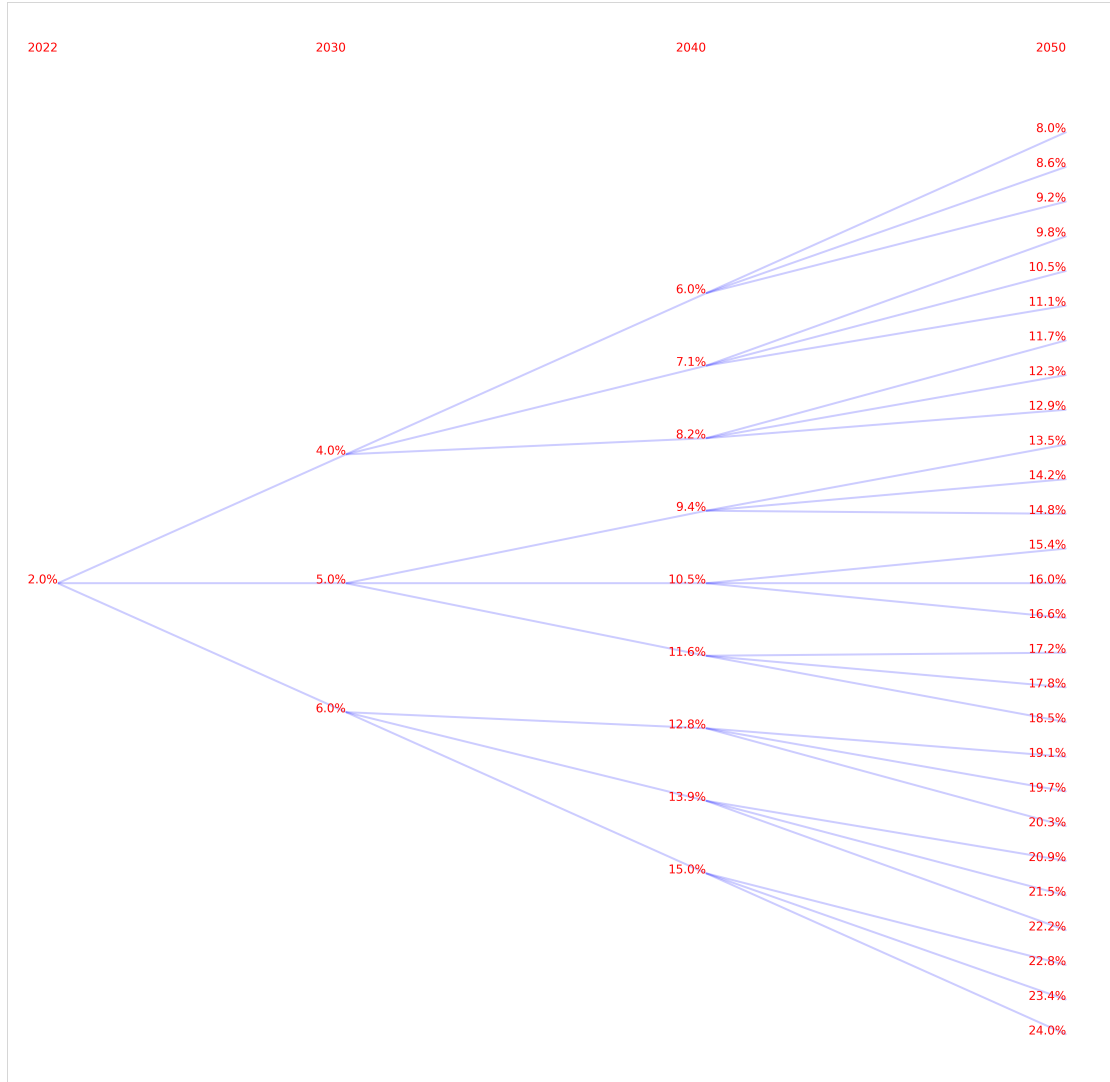


Figure 4.2: Scenario approach of the hydrogen demand in term of percentage of the total energy consumption

4.4 Hydrogen Production

Hydrogen is already used today in various sectors. However, it is mainly produced with natural gas (through a technology called Steam Methane Reforming) or even coal and other hydrocarbons. It is used in the fertilizer value chain, for example, as key ingredient to produce ammonia. The main problem emerging from this strategy is the emission of carbon dioxide.

In the solution presented, we are mainly interested on green hydrogen using

electricity from green sources. We focus then on different types of electrolysis.

Electrolysis is a technology that produces hydrogen and oxygen from using electric current and water ($2H_2O + Electricity \rightarrow 2H_2 + O_2 + heat$). This technology is already well developed and ready for large-scale industrial utilization. The most widespread typologies of electrolyzers are:

- Alkaline electrolyser (AEL) is the most mature. It uses electricity and water to produce hydrogen, and rejects only oxygen.
- Proton Exchange Membrane Electrolyser (PEMEL) is just a variation of the Alkaline technology but is more adapted for small capacities because it was historically used in that way. It produces slightly purer hydrogen.
- Solid Oxide Electrolyser (SOE) is not mature yet but seems very promising.

Our model is able to handle multiple technologies for production and transport, letting to the solver the choice of most judicious to use. However each technology added to the model comes at a certain computational cost. For these reasons our analysis will focus on the most promising of the already existing technologies, i.e. Alkaline Electrolyser. Its lifetime is around 10 years, i.e. the same time as our design time-steps. We can thus keep in mind that between time-steps, we can upgrade the technology to more efficient ones.

The cost of production of electrolysis is dominated by the cost of electricity. CAPEX costs and performance are well known and can be found in literature (see e.g. [26],[30]). We will use those estimations.

4.5 Hydrogen Transport

Hydrogen is a fuel, but its transport is still subject to a lot of discussions. Under standard condition of pressure and temperature, it has a lower density than other fuels in the market and it requires low temperature to be kept liquid which induces a certain energy loss. Different strategies are under investigation. It is for example possible to transform the hydrogen into another carrier to transport it more easily.

Methanol is one possible option. It is liquid in atmospheric conditions and benefits from already widespread infrastructures and a more advanced knowledge. Methane is yet another solution, widespread as well, and is produced using exothermic reaction aptly called methanation. Methane has a low boiling temperature and

could then be easier to be kept liquid. However, these two options share the same problem: they are not carbon free. Moreover, their potential impact on climate would then contradict the initial objective of the study.

Ammonia is another possibility that could allow an easier transportation. While it is carbon-free with a moderate boiling temperature, ammonia is highly toxic and has no direct usage. The transformation technologies are unfortunately not at maturity yet.

One last alternative would be to use Liquid Organic Hydrogen Carriers (LOHC). They are liquid at ambient conditions, cost effective, safe and non-toxic.

The transformation of one product into another means a loss of energy and a supplementary cost that should be addressed when comparisons are made. Following the literature in the field (see e.g. [25]), "for short import and distribution distances, compressed hydrogen (CGH2) pipeline - truck transport with cavern storage is the most feasible option". As our case study focuses on Belgium and that relatively small distances are used, those are thus the two technologies that we will consider. Long range maritime shipping will be used only for the node connecting our network to the international trade but we do not have to worry about this part now and we can consider that the hydrogen will be put in the same form of our network upon disembarkation.

Costs for the different transport means that we have selected can be found in the literature (see e.g. [11], with costs estimates in €/ (MW km)). Key values are presented in table 4.1. Moreover, once a pipeline is built, it makes no sense that one could just add small additional capacities. To address this constraint, any capacity addition will have to be higher than a predefined minimum value.

Technology	CAPEX [€/ (MW km)]	OPEX [€/MW]	Lifetime [yr]	δs_q^{min} [MW]
pipeline	$8.2 \cdot 10^3$	$1.6 \cdot 10^{-6}$	50	1.6
train	$6 \cdot 10^3$	$3.2 \cdot 10^{-7}$	10	1.6
truck	$13 \cdot 10^3$	$1.6 \cdot 10^{-5}$	40	$3.6 \cdot 10$

Table 4.1: Transportation costs, lifetimes and minimum additional capacities

4.6 Electricity Market

The main cost of operating electrolysis technology, as mentioned above, is the cost of the electricity required for the hydrogen production. For the case study, we are going to follow this logic and set the opex of maintaining the technology to zero in order to focus only on the electric costs. According to European statistics [10], the cost of electricity in Belgium in April 2022 was 0.14 €/kWh for non-household consumers. For the neighbouring countries, the value was : 0.14 €/kWh in Germany, 0.12 €/kWh in the Netherlands and 0.10 €/kWh in France.

Electricity prices for non-household consumers, second half 2021
(EUR per kWh)

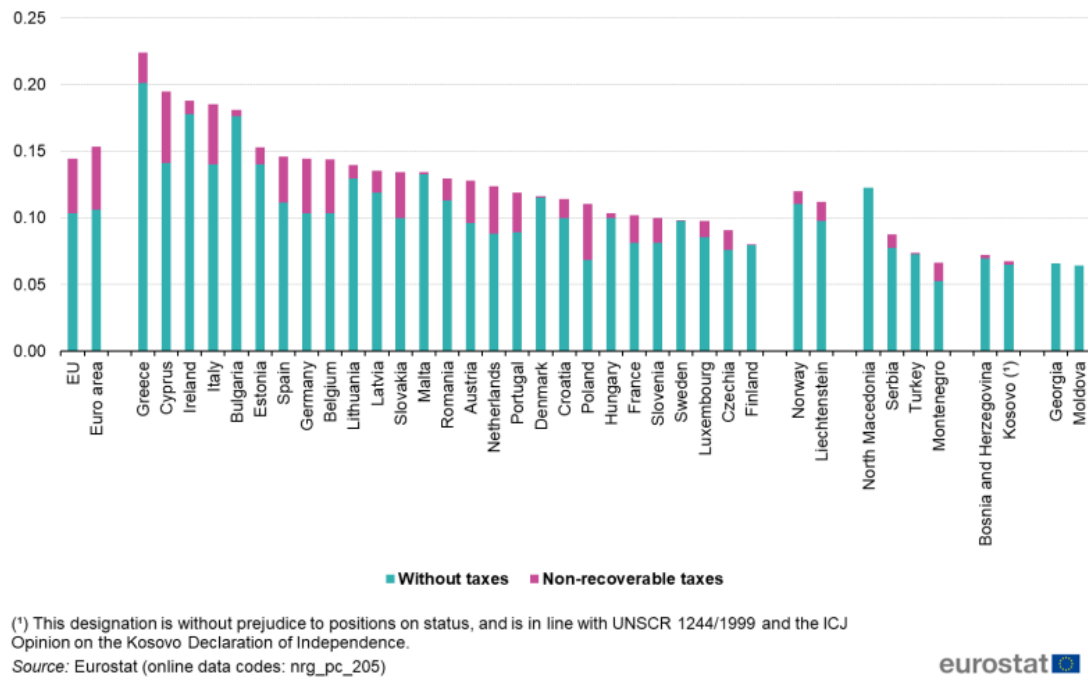


Figure 4.3: Electricity prices for non-household consumers in Europe [10]

While we have estimates of future values of hydrogen demand, we only possess current values for the prices to buy electricity on the market. To generate a tree of scenarios, we will use the same method than for the demand, with a set of equiprobable values between two bounds. Since we do not have information on the future electric market, we decided to take very high and very low value as bounds to highlight its impact on the solution.

We note as a reminder that we are only interested in the market of green electricity, given that we want to produce green hydrogen. As already highlighted in the introduction, not all regions have the same potential to generate green electricity. In our case, we have to take into account the availability of green electricity for all the nodes of our graphs. We will consider that we can import from a node a maximum amount of green electricity defined as the maximum possible production of green energy minus the consumption of the node in energy. It would indeed make no sense to use green energy for hydrogen production and not having enough green electricity for the local needs. By those computation we obtained the following numbers [11] : France (with a maximum green electric potential of 917.1 TWh), the Netherlands (142.2 TWh), Germany (200.3 TWh), Luxembourg (0.0 TWh) and Belgium (0.0 TWh).

However, these values are not a good representation of the green electricity available as regions of Belgium are deeply integrated in the European electric market. We will consider a 'copper plate' assumption which means that electricity is available at any node with the Belgian price of 0.14 €/kWh. This is the same value taken by Gabrielli on his case study of hydrogen for the Swiss network [11] which serves as basis for other parameters we use. We will discuss the impact of these assumptions in the Results chapter.

4.7 Hydrogen Import

In order to represent the uncertainty regarding hydrogen import prices, we use the same scenario approach as for the demand. In [9], hydrogen import costs are predicted to range between 65 and 90 €/MWh by 2030-2035 and potentially down to 55-75 €/MWh by 2050. These ranges are sometimes considered optimistic, hence we propose to use wider intervals in our scenarios, as illustrated in figure 4.4.

We note that the case study focuses on Belgium. As Belgium is a relatively small market, we will consider that the international market represents a potentially infinite source of hydrogen.

4.8 Trees of Scenarios - Uncertainty Combinations

To start with, we will work with two sources of uncertainty: hydrogen price and demand. We will not consider that their variations are directly correlated. For each of the two values we consider three possible next states. By combining the

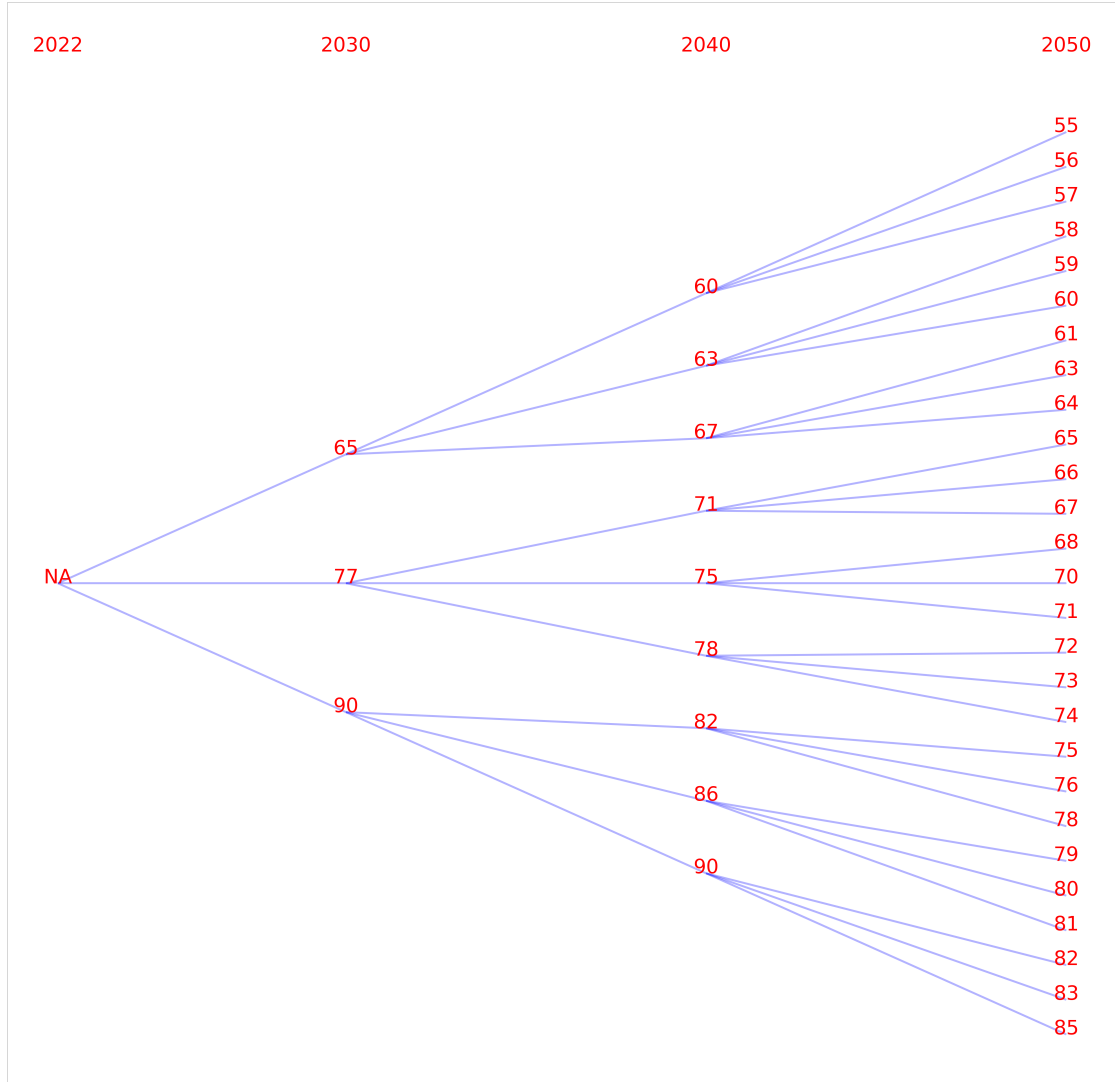


Figure 4.4: Scenario approach of the import cost of hydrogen [€/MWh]

two, we get 9 possible child states. In this example, the price of electricity is kept constant and only the price and demand of hydrogen are uncertain.

In a second stage, we will use all three uncertainty parameters: hydrogen price, demand and electricity price. For each of those we will consider 2 scenarios which make 8 possible scenarios in total. The combinations of values are given in figure 4.5.

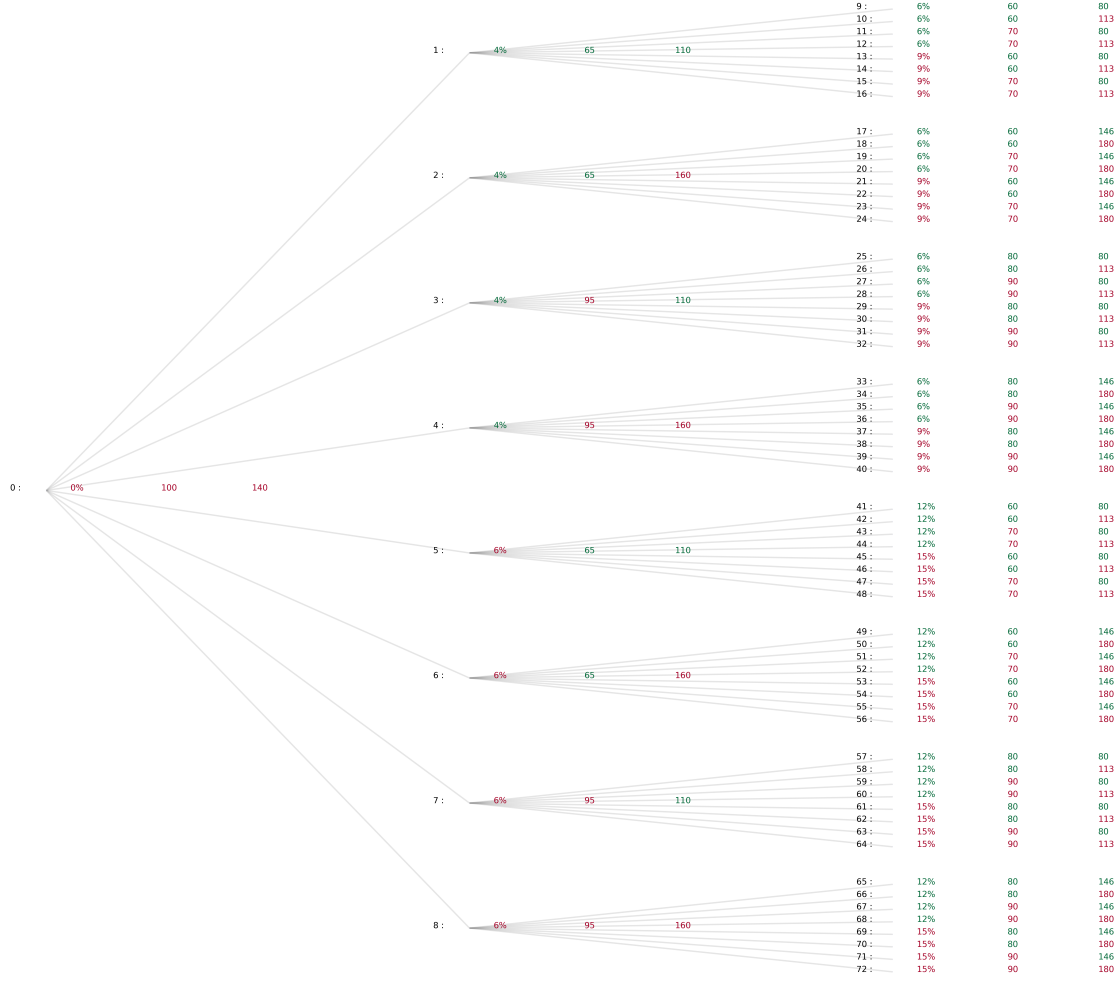


Figure 4.5: Scenario approach for all the uncertain parameters with their values along the state tree in the first three periods

4.9 Implementation Choices

In this section we present the design structure of our optimization model. Broadly speaking, the problem is divided in three blocks. The interactions between the different blocks is illustrated in figure 4.6.

First, as explained in the previous section, trees of possible scenarios are generated by our first block. These trees are then submitted to the 2-level iterative algorithm, to find optimal solutions. A first part of the decision variables are

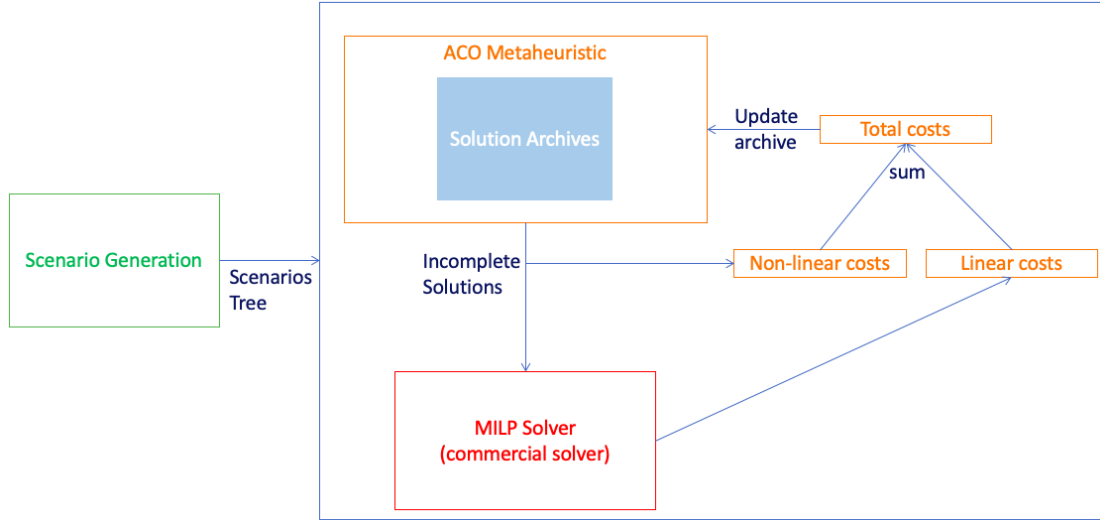


Figure 4.6: Meta-heuristic structure

handled by the Ant-Colony meta-heuristic (second block or "upper" level). At each iteration, the algorithm creates a certain number k of incomplete solutions, each containing values for these decision variables. The other part of the solution is build by a classic solver (third block or "lower" level), using the decision variables provided by the master as constants. The two solutions combined give the total corresponding cost, that is used by the Ant-Colony algorithm to update and rank its archive of best solutions evaluated this far. The archive is then used to create the new solutions for the next iteration.

The solution we have build is based on an existing "upper" level, i.e. a generic Ant-Colony meta-heuristic aimed to solve a range of supply chain planning problems. Specifically, the contribution of this master thesis was to develop the first and third blocks to solve the hydrogen case study: generate feasible trees of scenarios taking all dimensions of the problem into account (first block), develop a fit-for-purpose solver the Ant-Colony algorithm can call to generate optimal solutions (third block).

The code developed is open and generic. It allows to handle different types of variables (linear / not, discrete / continuous), constraints and objective functions. The Ant-Colony meta-heuristic can be turned on and off in function of the complexity of the problem to solve and the execution time required. Different solvers can also be used in the "lower" level. We have chosen to use the commercial solver **gurobi**, for our development.

In the next chapter, we will analyze the results of our experimentation using this approach and algorithm.

Chapter 5

Results

In the previous chapters, we have defined with precision our problem, our model, the case study and the related implementation design choices including the algorithm used. In this chapter, we present the results of our approach on the case study.

We start with a short introduction to the experimentation plan and the way we propose to represent solutions graphically. We then assess solutions found for a simplified version of the model and 4 families of variations: key constraints are added to the base case and their respective impact on solutions are discussed. We follow with an assessment of the time complexity of the algorithm. Finally, we discuss the approach: its benefits but also its limitations, should we want to tackle increasingly more complex (real) problems.

5.1 Experimentation Plan and Graphical Representation of Solutions

Base case and variations. In the following sections, we will analyze the results of our approach on a base case and 4 families of variations of constraints / parameters:

- Base case: In this first case we allow the algorithm to create production capacities of any size including very small ones, i.e. kW by kW , and we also do not penalize the network for light loading of installed capacities
- Variation 1: We assess the impact of adding constraints on production capacities and utilization factors
- Variation 2: We look at the effect of uncertainty on the cost of electricity imports

- Variation 3: We analyze the implications of shifting from uniform to biased probability distributions for some parameters such as demand scenarios
- Variation 4 : We look at the effect of allowing broader ranges of values for the hydrogen import cost.

Trees of states and scenarios. A solution returned by our algorithm is an optimal design of the network for all possible states of the system. It is important to remind that our approach is multi-stage, multi-scenario. This means that each state corresponds to a specific design of the network. Each of these states results from the decisions made during the previous states / past time-steps, e.g. the network in 2050 results from decisions made at time-steps 2020, 2030 and 2040. It also means that the solution should at each state find an optimal design to serve a future demand, starting from an already existing network, resulting from past decisions, and using the new information available at that time. The structure of the multistage allows us to model this progressive information flow.

Let's consider the following example: At time 1, we do not know anything except that the demand should range between 4% and 6% in time 2 and between 6% and 15% in time 3. If we are in the case where the demand is 4% after the first time-step, then we know that the demand in the second time-step will be between 6% and 10.5%. On the contrary, if the demand in time 2 is 6%, then it will range between 10.5% and 15% in time 3. The values of the parameters follow a certain trend as time progresses.

For this reason, even if the two states start with the same network at time 1, the decision made in the state of time 2 are not the same since they have to prepare for different demand thresholds that will arise in time 3. This observation holds for all uncertainty parameters and not just for the hydrogen demand.

It is important to note that this is a key design choice of our model. Indeed, making all scenarios for a next time-step possible regardless of the current state would make the multi-stage useless: there would be only one network design by time-step since the uncertainty would not decrease as time progresses. The situation would indeed be the same for each state of the same period since they would have the same base network design and expectation for the future.

In our general model, the main decisions are the size and location of the production capacities as well as the pipeline and other transport modes to install. The tree in figure 5.1 represents how the states succeed to each other and shows the

evolution of the uncertainty parameters.

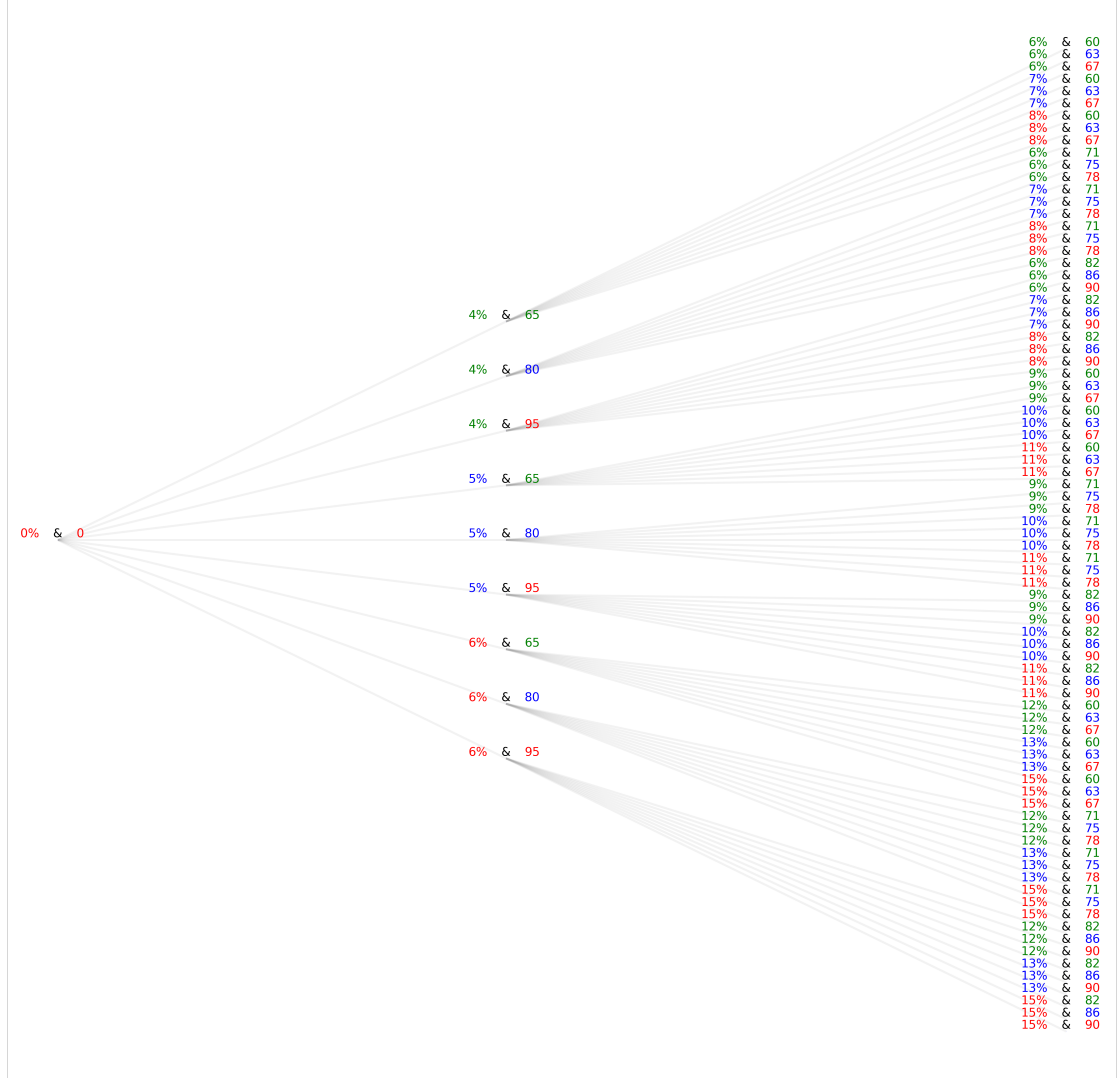


Figure 5.1: Import cost and demand of hydrogen by state, deployment of the uncertainty along the state tree

The first set of important decisions will be made at time 1 and the resulting design will thus be the same for all the states at time 2. However, when arriving at those states we gain new knowledge that the solution will use to upgrade the network adequately. In time 3, there are thus as many different designs as the number of states in time 2.

Graphical representations of solutions. In order to present the solutions we will draw the resulting networks at different key states, with the installed capacities of transport and production. The nodes of the graph are the different regions of Belgium, a node for the North Sea import, and also one node for each neighbouring country. The edges represent the installed pipelines and other transport modes to link those regions. The size of each element in the graph is proportional to its size in term of kW . Thus more capacity means bigger nodes and thicker edges. For the values usually given in kWh/yr we divide them by the time in one year to compare them with the capacities that are already given in kW , as if the network was working without interruption at constant rate.

To ease the readability of the graphs, the following color codes will be observed : Blue correspond to the production of hydrogen, Green to import of hydrogen, Red to exports, and Orange to demand.

We will see that the capacities installed in the solutions are not always used at their maximum potential. Indeed, the actual production and transport load follow the demand at a particular state and it is common that certain capacities are installed to handle the scenarios with higher demand. The capacity loading of the pipeline is noted as percentage directly on the graph of the network. For the production capacity, finally, the percentage is given next to the id of the node.

5.2 Base Case: Unconstrained Capacity and Utilization

Our first experiment is a theoretical case in which we allow the algorithm to build "granular" production capacity as required, implicitly assuming no scale effects, and to under-utilize capacities once installed, assuming that most costs would be variable. Whereas these two aspects do not match the economic reality, the experiment allows us to test the algorithm in simple conditions.

In figure 5.2, one can see the resulting network solution of the algorithm at time 2 (2030). In this solution the bigger nodes with high demand are kept isolated since they face a high enough demand and they are too far from the imported hydrogen source NS. In the case of lower demand, the production facilities are used only at 66%. On the contrary, the smaller nodes (BE21, BE32, BE33) are closer to the source of imported hydrogen which makes this option cheaper in comparison to local production.

State 1 (t = 2030) Investment : 40.515Md€
Hydrogen demand : 4% | Hydrogen price : 65.0€/MWh | Electricity price : 140.0€/MWh

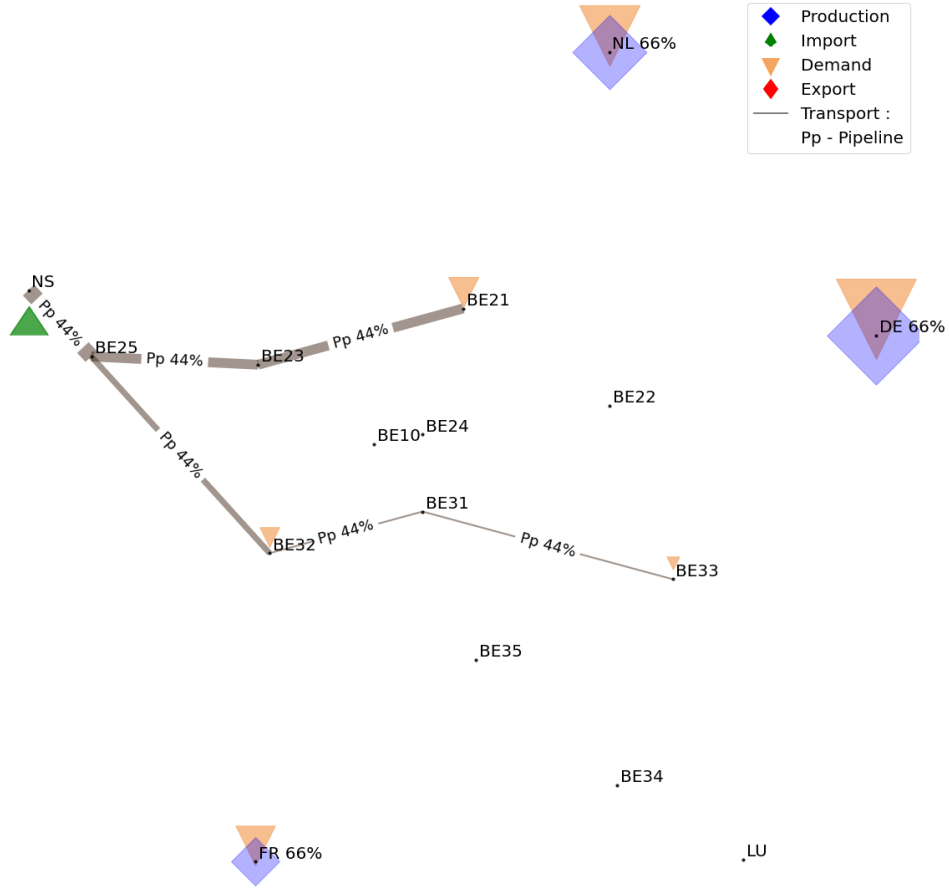


Figure 5.2: Base case: Solution network design at time 2 (2030) + capacity utilisation for state 1 (Low demand and Low cost)

5.3 Variation 1: Capacity Concentration and Utilization Factor

We will first try to add specific values for parameters `minBuildCapacity` and `minload`. The former one will constrain the way we build production capacities. It indeed makes more sense to build electrolyzers of a certain minimum size vs kilowatt by kilowatt. Even if it is technically possible to develop smaller production sites, economies of scale make the installation cost non-linear. Note however that for higher capacity values, the cost is quasi-linear. Working with a capacity threshold is aligned with our decision to linearize the CAPEX. Lower numbers would not have been used by the solution anyway since they are less efficient.

When new capacities are installed, they thus have to reach a certain amount defined by the parameter `minBuildCapacity`. This will create a reaction of the model linked to another important parameter called `minload`. With this one, we constrain our facilities to be used at least at `minload%` of their maximal capacity. In our example, the demand after the first time-steps is not high enough. If we build at least `minBuildCapacity` capacity at each node presenting a demand, our production surpasses the needs and the network is forced to export the surplus with a negative financial impact (given unbalance between total production cost and export price). In this case the solution starts to concentrate the production in certain regions, those with higher demand, and use the transport technology to connect those producing nodes to the other ones.

To illustrate this, we run the algorithm with the following values : `minBuildCapacity` is set to 10, `minload` is set to 80%.

The solution is different from the base case as all the nodes are now indirectly connected as shown in figure 5.3. Pipelines are built but only to be used in the case of lower demand. They are connecting NL, DE and FR to the rest of the network. This could seem unsettling. In fact, the capacity installed depends on the high demand case. The load then reaches 100%. However, in the lower demand scenarios, the production is so high because of the 80% `minLoad`, that it has to supply the demand of BE21, BE32 and BE33 to get rid of its surplus. In the same way, other pipelines are used for high demand scenarios. Indeed, in those scenarios the production of neighbouring countries is kept to handle local demand and the Belgian regions have to rely on the North Sea import. In order to handle those two situations, the algorithm builds the pipelines for the two options at time 1, even at the risk of never using them. The solution is bimodal for each Belgian node with the product coming from one side for low demand scenarios and from another side in higher demand.

Note that the variation of the demand is expressed in percentage of the total demand. This means that nodes with higher absolute demand suffer from higher variations. Node DE can supply itself in the worst case. However, it produces way too much with the 80% minimum loading we fixed. In addition to the new pipeline to BE33, hence the solution given by the algorithm exports part of the surplus. It will not share it with the rest of the nodes that have to be supplied using other ways. This suggests that the cost required to allow the network to efficiently dispatch this overproduction is too high compared to the loss caused by exporting hydrogen.

State 1 (t = 2030) Investment : 47.242Md€
Hydrogen demand : 4% | Hydrogen price : 65.0€/MWh | Electricity price : 140.0€/MWh

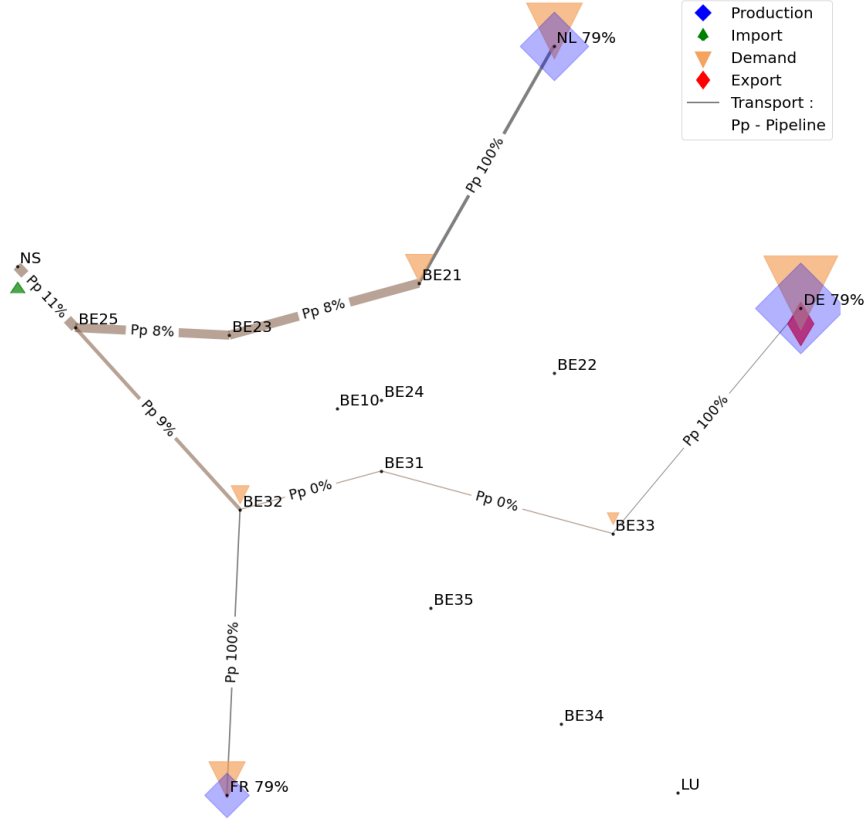


Figure 5.3: Variation 1: Solution network design at time 2 (2030) + capacity utilisation for state 1 (Low demand and Low cost) with $\text{minBuildCapacity} = 10$ and $\text{minLoad} = 80\%$

We now apply the same strategy for 2030 to 2040 and for 2040 to 2050. The design is the same even if the numbers are not the same as the network has already been through two adaption phases after its initial design. The final network (2050) is given in figure 5.4 - high demand case where we see that the situation in time 4 is similar to time 1, with certain pipelines not loaded at all.

Figure 5.5 shows the low demand case in time 4, too, with a similar structure

State 84 (t = 2050) Investment : 235.364Md€
Hydrogen demand : 24% | Hydrogen price : 75.0€/MWh | Electricity price : 140.0€/MWh

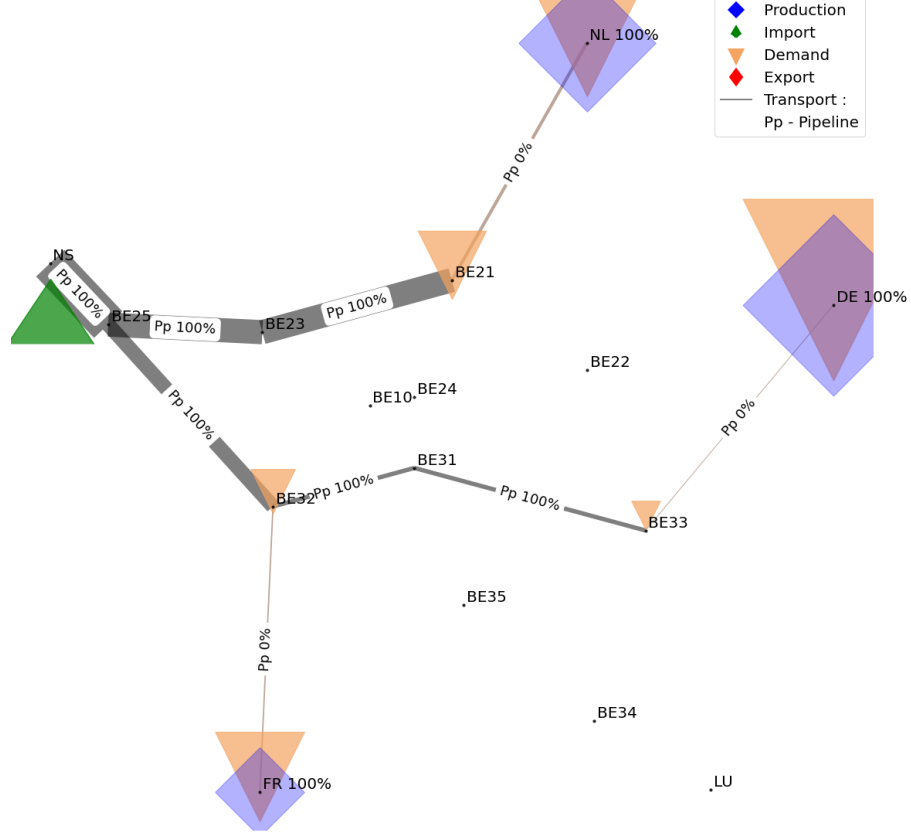


Figure 5.4: Variation 1: Resulting network design at time 4 (2050) + capacity utilisation for state 84 (Highest demand and Highest cost) with `minBuildCapacity` = 10 and `minLoad` = 80%

of solution.

The two uncertainty parameters we worked with to produce the previous results were the hydrogen demand and the hydrogen cost. If we plot the distribution of the final cost for all possible scenarios in 2050, the total cost follows the same trend as the demand (see figures 5.6 and 5.7). Indeed, most of the cost in this example comes from imports of product and because those are chosen after solving sources

State 21 (t = 2050) Investment : 79.452Md€
Hydrogen demand : 8% | Hydrogen price : 55.0€/MWh | Electricity price : 140.0€/MWh

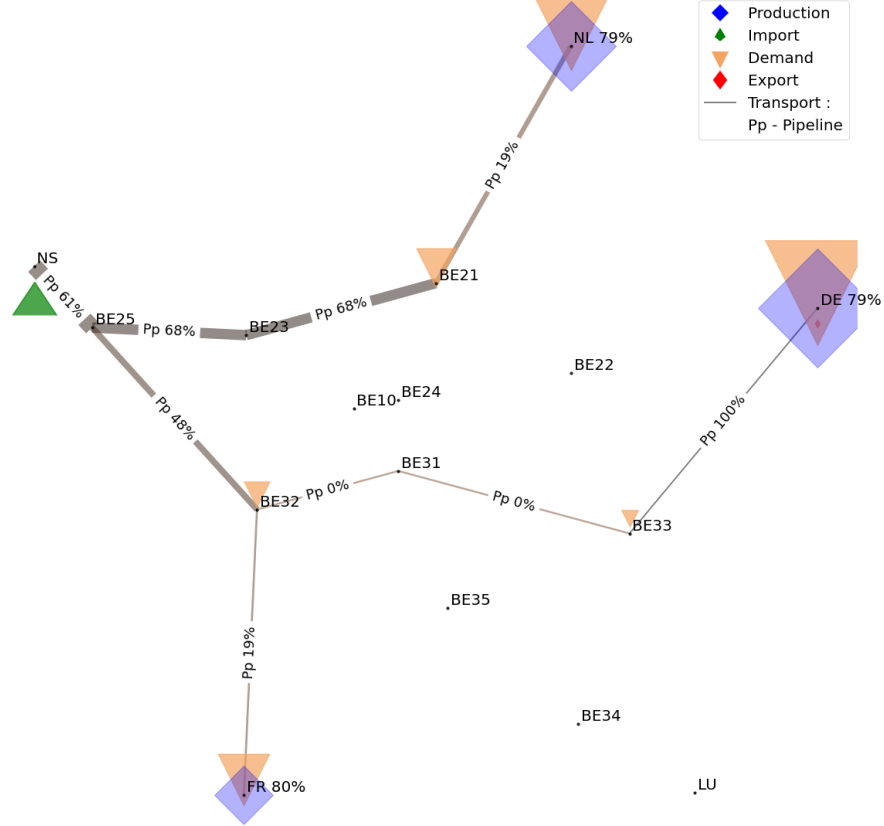


Figure 5.5: Variation 1: Resulting network design at time 4 (2050) + capacity utilisation for state 21 (Lowest demand and Lowest cost) with `minBuildCapacity` = 10 and `minLoad` = 80%

of uncertainty the cost will generally be adapted.

Moreover we can see in figure 5.8 that most of the production comes from local capacities. Thus most of the cost comes from imported electricity which explains the little influence of the hydrogen import cost variation. Note that the quantity of imported hydrogen increases for higher state (those in time 3 and 4) simply because the total demand increases. One could be tempted to advance the price of

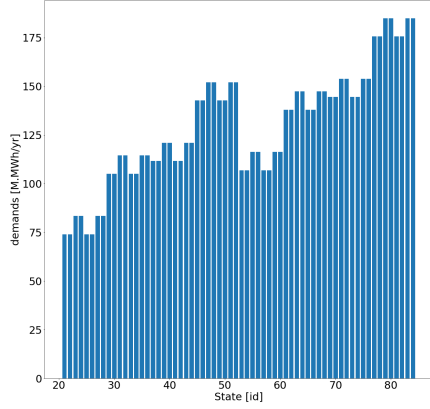


Figure 5.6: Cumulative demand served from 2020 to 2050 for each possible sequence of scenarios with `minBuildCapacity` = 10 and `minLoad` = 80%

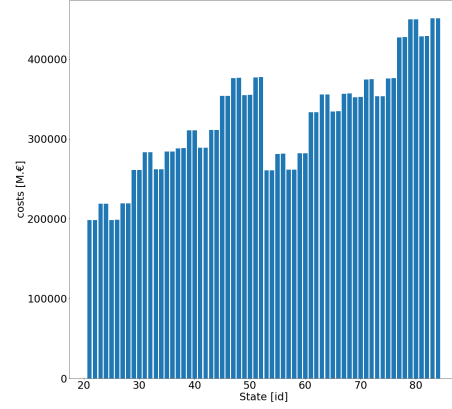


Figure 5.7: Cumulative cost from 2020 to 2050 for each possible sequence of scenarios with `minBuildCapacity` = 10 and `minLoad` = 80%

hydrogen will slowly decrease while the cost of electricity could remain the same or even increase. The cost of one unit of imported hydrogen at a particular node would be the cost of import plus the cost of transport for this unit from NS. If imported hydrogen would really become cheaper than local production, the switch would be total as transport costs are linear with capacity. All hydrogen for that node would end up being imported.

The expected total cost goes from 2,157 Md € for the base case to 3,208 Md € for this example. The increase in total cost between the two situations is 1,051 Md €. This is thus a non-negligible amount since this is a strong constraint to impose to our network.

5.4 Variation 2: Uncertainty in Cost of Electricity Import

Until now, we have only considered uncertainty on hydrogen demand and hydrogen cost. There was however in our case study another uncertainty source, namely the cost of buying the electricity necessary for the hydrogen production in our

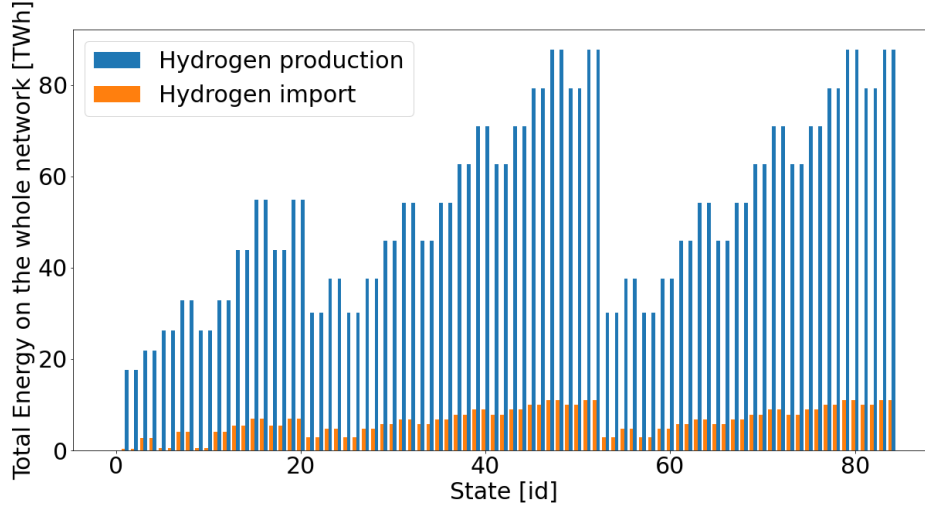


Figure 5.8: Comparison between the amount of hydrogen that is imported vs produced directly on the network

production facilities.

From the tests conducted with the three uncertain variables at the same time (hydrogen costs and demands as well as electricity costs 4.5), it seems that the algorithm still finds the same network design solutions. The smaller nodes are connected both to the North Sea and to the neighbouring countries and receive mainly from one of the two paths depending on particular demand realisation scenarios.

The system adapts mostly to the demand more than the cost. However, in the "Variation 4" section below, we will trigger a response of the algorithm by increasing the relative difference of costs between imported and produced hydrogen, by playing on the costs of hydrogen.

5.5 Variation 3: Biased Probability Distribution

Until now we have always considered that all scenarios had the same probability. In the previous chapter we explained that the possible values will be generated using discrete 'steps' or ranges. For example, we consider based on the literature that hydrogen demand should range between 8% and 24% in 2050. As shown before, upper

and lower bounds have more influence on the resulting solutions than intermediate values, hence generating numerous scenarios between them is not very insightful.

Here, we are modifying the probabilities of scenario. For instance, we define a set of scenarios with very biased probabilities: the demand will be low only in 10% of the cases and at its highest in the remaining 90%.

In the solution found by the algorithm (see figure 5.9)), the nodes FR and DE are not linked to the rest of the network like before (see figure 5.3). Those were indeed connected to the rest of the network to share their surplus of production when the demand ended up being low. Now the case of low demand is much less likely, the loss caused by the surplus of production has less impact on the objective function and the introduction of new pipelines to handle these events is not worth anymore. Indeed pipelines have to be build in time 1 and since they are useless when demand is high, they are useless in 90% of the case.

5.6 Variation 4 : Large Variation in Hydrogen Import Costs

As observed in the previous examples, the imported hydrogen is cost-competitive compared to local production. We did however base those runs on the most optimistic forecast for the hydrogen costs (5.1). In this last example we will consider a case with larger variations. We will also remove completely the constraint on production loading.

The solution starts with the same design as in time-step 2 of the base case solution. It will continue on the same path when the hydrogen cost follows the prevision and the price continues to decrease. On the contrary when the hydrogen price is at its peak at time 2, the network will undergo new developments.

The results show that new production capacities have been installed in BE25 and BE21. However, when the price decreases between period 2 and 3 (see figure 5.10), the network will decide to continue to rely on import, letting unused those new production capacities. On the contrary, when the price of hydrogen continues to increase and reaches new heights in state 12, then import is completely abandoned to rely only on local production since it is now the cheapest option (see figure 5.11).

This choice of node BE25 as new production facility is not trivial. The network is

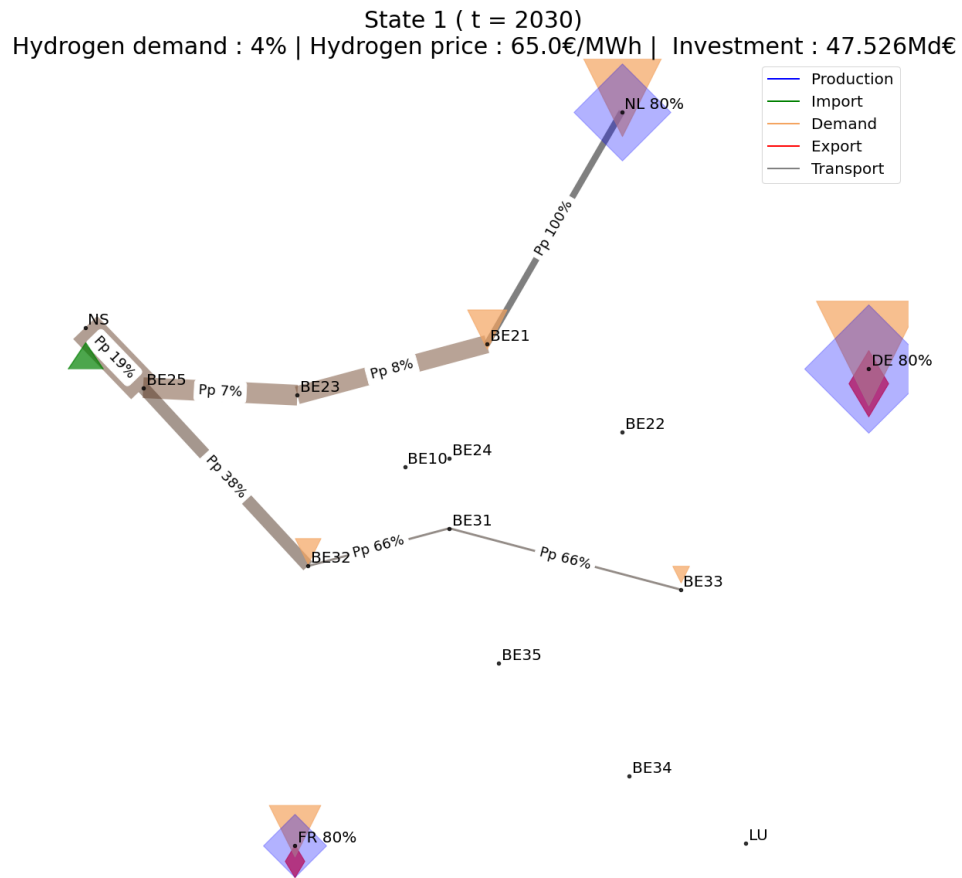


Figure 5.9: Variation 3: Solution network design at time 2 (2030) + capacity utilisation for state 1 (Low demand and Low cost) with biased probability

already organised around this node as it is a forced passage for imported hydrogen coming from NS to enter the rest of the network. The production in BE25 can then be distributed to the rest of the network using the same pipelines.

State 9 (t = 2040) Investment : 62.827Md€
Hydrogen demand : 6% | Hydrogen price : 220.0€/MWh | Electricity price : 140.0€/MWh

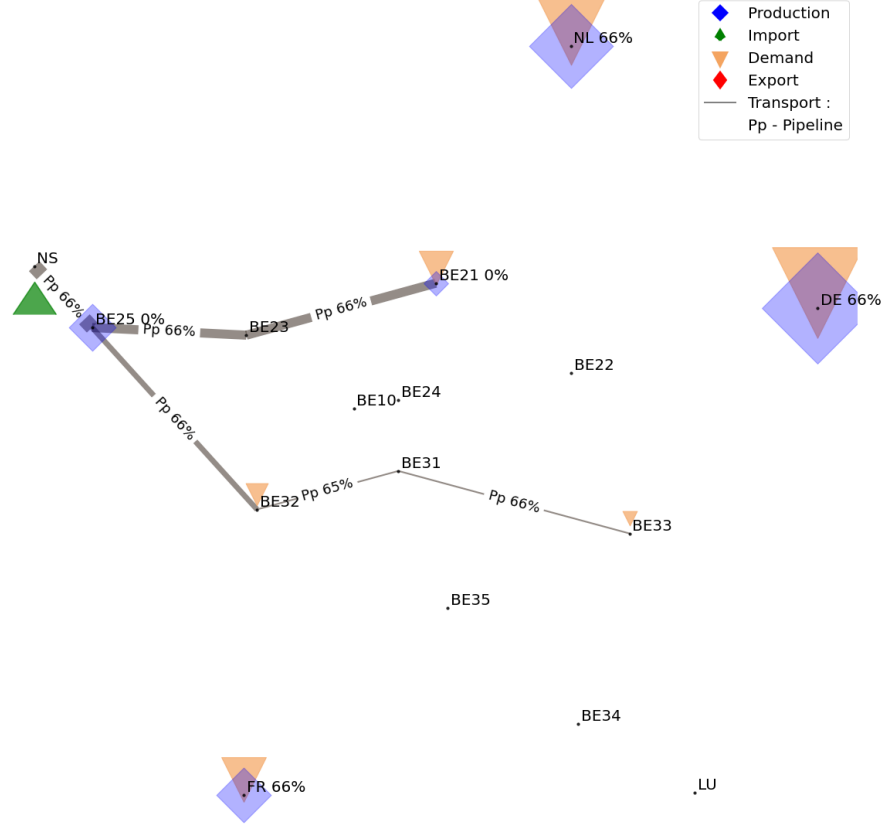


Figure 5.10: Variation 4: Resulting network design at time 3 (2040) + capacity utilisation for state 9 (Low demand and Low cost) with large variation in hydrogen import costs

5.7 Complexity Analysis

In this short section, we discuss the time complexity of our algorithm.

State 12 (t = 2040) Investment : 95.848Md€
Hydrogen demand : 9% | Hydrogen price : 300.0€/MWh | Electricity price : 140.0€/MWh

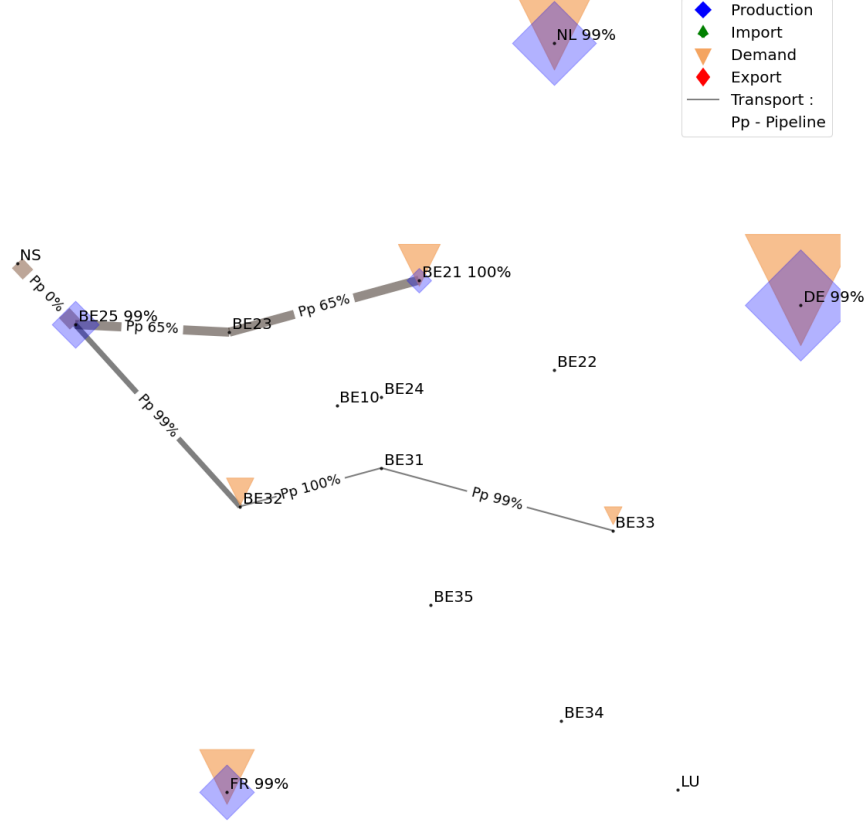


Figure 5.11: Variation 4: Resulting network design at time 3 (2040)+ capacity utilisation for state 12 (High demand and High cost) with large variation in hydrogen import costs

A Impact of Number of Uncertainty Parameters

In the previous section we have analyzed a base case model and variations of our constraints and parameters. Clearly, adding multiple scenarios together leads to an exponential increase of the number of states to consider with clear implications on the total time to find optimal solutions.

The number of possible scenarios depends on the number of uncertainty parameters u and the number of possible scenarios for each parameter v . The total number of scenarios is thus equal to the number of possible combinations of future realisation for all parameters: v^u .

At time 1 corresponds only one state, which is the initial state. There is no uncertainty there yet. At the first time-step, we have a number of states equal to the number of evolution scenarios s_c . Each of these states will also give a number equal to the number of evolution scenarios in the second time-step. The total number of state is thus given by:

$$|\mathcal{S}| = \sum_{t \in [1, |H|]} s_c^{t-1} = \sum_{t \in [1, |H|]} (v^u)^{t-1}.$$

As shown in figure 5.12, introducing a third uncertainty parameters already drastically increases the number of states.

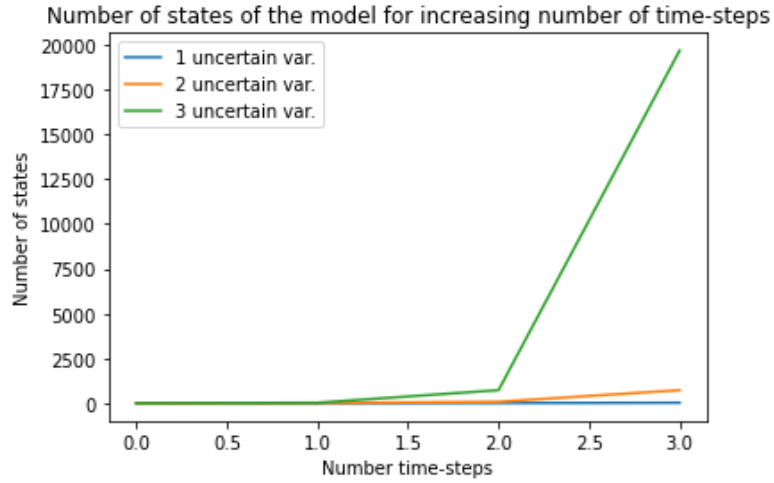


Figure 5.12: Comparison of the complexity in term of number of state of the model for different number of uncertain variable

B Linear Case (MILP)

To solve the problem we constructed a MILP model that is solved using the `gurobi` solver. The solver uses the simplex heuristic. When the algorithm does not find a solution in the first 200 sec, it will usually continue to run without significant improvement for potentially long time before converging. For easier two-stage cases,

$ H $ = Nb. Periods	2	3	4
Nb. States	9	73	585
Nb. Continuous Variable	4791	38839	311223
Nb. Integer Variable (binary)	1107	8979	71955
Presolve Nb. Continuous Variable	874	7866	63802
Presolve Nb. Integer Variable (binary)	56	504	4088
Presolve Nb. Constraints	66	594	4818
Solving time [sec]	~ 1	~ 1	76

Table 5.1: Evolution of the complexity of the problem with the number of periods H

the algorithm takes only a couple of seconds.

The time complexity to find a solution is mainly dependent on the problem complexity more than its size even if the size can be a good proxy. Indeed, certain parameters can lead to large variations of computing time for different input values.

In order to gain more insight on the evolution of the problem complexity as the number of states increases, table 5.1 shows the evolution of complexity for 2 possible outcomes and 3 variables.

Because of the problem's structure, the complexity of the model increases exponentially with the number of periods $|H|$. The model has at each state a whole set of decisions to make. The number of states depends on both the number of scenarios and the number of time-steps we want to consider.

C Non-Linear Case (MINLP)

At each iteration of the meta-heuristic solver, a number k of new solutions are built using the solution archive. For each of those k incomplete solutions, their 'quality' needs to be assessed. We do that by calling a 'slave' algorithm inspired from the initial linear version, which will solve the same problem but with some variables replaced by parameters. Those are the variables handled by the meta-heuristic and for which values are contained in the archive.

The increase of complexity of the problem has two implications on the execution time. When the number of states in the problem increases the complexity of the slave problem will also increase (see figure 5.13). More states also means more

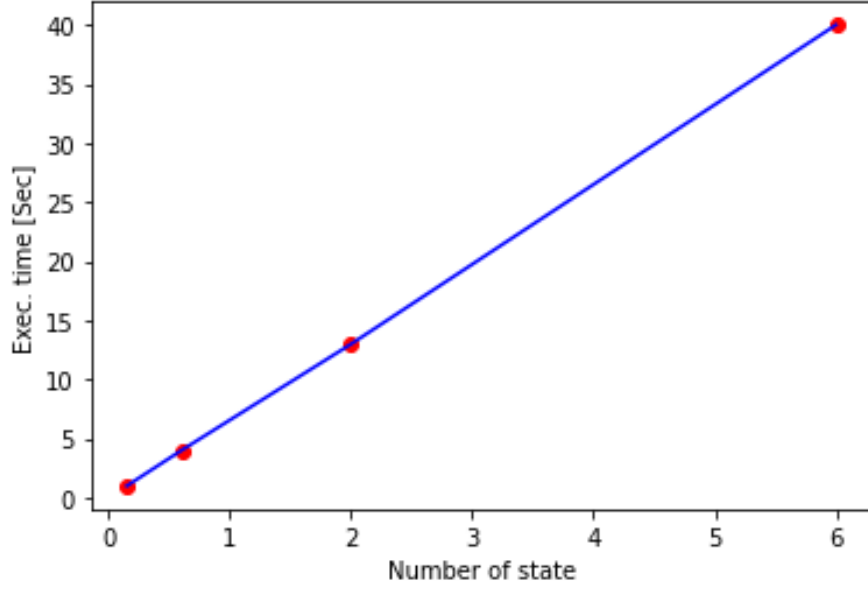


Figure 5.13: Evolution of the execution time for the slave algorithm with increasing number of state in the model

variables and this number will have an impact on the convergence speed of the main loop. Indeed the number of iterations to have the same precision will be higher if there are more variables to tune by the algorithm.

The run-time of the meta-heuristic is mainly determined by those calls to the slave algorithm. We can thus compute how the execution time will increase as the number of states of the problem increases:

$$T_{exec} = k_{ants} \cdot t_{exec} \cdot it.$$

For a classic case with around 100 ants and 40 states, the formula gives $T = 100 \cdot 73 \cdot 1000 = 7.3 \cdot 10^6 sec$ for 1000 iterations. This correspond to 2027 hours or 84 days.

D Trade-Off: Number of Scenarios vs Precision

As we have seen, the complexity of the model is rapidly very high due to the multi-stage approach. However, this execution complexity does not always come with significant improvement of the solution.

From our experiments, it turns out that the results using larger scenario trees with more possible scenarios for each uncertainty parameter are sensibly the same.

Most of the time, increasing the precision with the number of scenario does not create any real variation in the solution. The optimal solution found is mainly influenced by the upper and lower bounds and thus the two extreme scenarios. The decisions are mostly guided by the constraints and if the network can handle those scenarios, it can handle any scenario in between, without exploring all combinations. Our model definition allows a certain flexibility in how we use installed capacities.

For this reason, we used simplified scenario trees to produce the results of sections 5.4 to 5.6: only 2 branches at each state instead of 3 for each variable. This indeed drastically reduces the time complexity and allows us to work on problem with more periods, while keeping solutions very close to the ones with more scenarios.

5.8 Assessment of Our Approach and Results

The algorithm developed finds optimal solutions for the base case as well as for a number of variations we introduced to make the model closer to real cases. The basic objective of our work has thus been achieved. However, the model suffers from limitations we are covering in this section, and could inspire future work.

First our model is of "brute-force" type, as we create a potentially large set of scenarios and calculate an optimal result by assessing all scenarios at the same time. This method is computationally heavy, as we have seen. On the one hand, the models we were able to run were not yet realistic enough - a number of constraints and parameters could be added to get closer to reality, at the expense of an explosion of complexity and execution time. On the other hand, this execution time is not a severe issue as our goal is to define a long-term industrial plan and not to optimize a network with real-time changes of constraint and parameters; even a couple of hours / days of calculation time might still be acceptable in this context.

Second, the two main sources of costs in our model are the CAPEX and the cost of importing the material, where the latter largely surpasses the former. However we defined the decision of import as second stage variable since we wanted to keep the balance constraint tight. This is the reason why the solution is quite simple in the first base case example. However there are links between the imports of carriers and some decisions taken thanks to parameters such as `minLoad`. Indeed this forces the model to use our capacity and to force an important cost if the network was not well designed. What stands out is the tendency for the algorithm to dodge any surplus of production that would force him to export.

Third, the current algorithm does not take into account the gain of remaining production and transport capacity at the end of the simulation. All production and transport elements have a defined lifetime. When we minimize the cost, we do it with the CAPEX, the OPEX and the import costs at the same time. The total cost we minimize is only over a predefined set of time periods. However the real consumption and usage of the network could potentially continue after the last time-step of the simulation. If those additional time-steps were taken into account, the optimal decision could be different, with more investment into long-lifetime, potentially more expensive technologies that could continue to serve in the future.

Fourth, a big advantage of working with multi-stage is to see the solutions adapt to the scenarios as time progresses and uncertainty unveils at each successive state. We observe that the model indeed adapts to new demand estimations and create robust final networks to face diverse scenario trees. However, the solution given by the algorithm does not present this particularity for some uncertainty parameters such as cost values. It seems indeed that their variation ranges have to be very large to create new solutions. The objective function could be modified to give a higher weight to even small cost changes, should this dimensions become more important.

Finally, as partly explained here above already, at this stage, our algorithm does not add significant precision to the solutions generated by adding time-steps or multiplying scenarios. To add significant precision, and thereby generating closer to real-life solutions in reasonable time, we could optimize our algorithm or even revise the underlying implementation concept itself.

Chapter 6

Conclusion

In a context of acceleration of efforts to decarbonize economies across the world, and assuming that hydrogen will play a central role in the next decades, this master thesis aimed to develop an approach to optimize the future hydrogen production-distribution system of a geographic region, while taking into account major sources of uncertainty in the long-term evolution of the system, and to apply / test this development on a case study: Belgium.

Hydrogen Supply Chain models have steadily grown in complexity over the years. The models tend to incorporate more and more the different elements of the value chain from production to distribution. However, the uncertainty inherent to this type of problem, was too often overlooked or simplified to its most optimistic behavior. Recent work on the subject has pointed to metaheuristics as possible solutions to more effectively address these complex models.

More specifically, we wanted to define optimal hydrogen production-distribution networks for the next decades (2030, 2040, 2050), taking into account probability distributions for key parameters such as hydrogen demand but also hydrogen import/export or electricity prices.

To get there, we leveraged a two-level algorithm with a meta-heuristic (here, the Ant-Colony) embedding classic solvers, the "upper" layer taking care of part of the variables while the "lower" layer focusing on other variables.

The solutions produced by our model, show that the most optimistic situations would offer very competitive prices for imported hydrogen compare to local production. However, relying uniquely on this solution could potentially backfire in worst cases. Depending on certain key parameters that regulate production facilities, the cost of security and robustness could be significant. Finally, the best locations for

local production would match the locations of import sources, to allow an efficient switch of the network from one type of product to the other.

With our approach, a solution of the problem contains all decisions made at all times and for all scenarios of our state tree. Our solution contains thus a huge number of variables compared to the deterministic case. Moreover, in the mathematical space in which our solution is defined, most of the values do not respect the constraints of our problem. Among the solutions produced by the Ant-Colony, most turn out to be unfeasible. It means that it is extremely difficult (and long) for the Ant-Colony to navigate in the space of solutions. Since most solutions are unfeasible, the algorithm iterates during a long time without progressing enough. This is true even for smaller problems where the simple MILP solver can find a solution more efficiently. Let's not forget however that we are interested here in long-term planning and that one solution only needs to be found when we want to build new capacity (we are not trying to solve a real-time problem). Under the assumption of our model this would be every ten year. Using the metaheuristic is fully justified for complex problems that would be practically unsolvable with the classic MILP solver, whereas it is less so in our (simple) case study.

Based on our work, we see a number of paths one could explore to enhance the approach followed.

First, as the chosen strategy creates a space of solutions that is difficult to navigate for the Ant-Colony, one could enhance the way potential solutions are generated outside of the metaheuristic itself, to save iteration time. In complement, one could also optimize the search space / data structures / code of the Ant-Colony to decrease its execution complexity.

A Second option could be to test or revisit other families of evolutionary algorithms, such as genetic algorithms or simulated annealing. Several research papers exploit genetic algorithms, for example, with positive results.

A third option would be to modify (simplify) the solution space to better apply to the metaheuristic capabilities. Several improvements could indeed be made here. Infrastructure usually do not stay fixed during long period and will be adapted in regard of the evolution of the system. The choice of the multi-stage allows to model this. It takes into account that different decisions will be made at different time periods in the future, to adapt progressively to the information arrival flow. It is however harder to handle than a simpler multi-period which corresponds more to the metaheuristic.

A fourth option would be work on the model of the problem itself. We have indeed presented that the metaheuristic struggles as most of the solutions in the search space are unfeasible. One would gain to express more of the problem constraints as objectives to minimize.

The current solution considers imports and exports as "second stage" variables since they are decided after the revelation of the uncertainty parameters. This allows the model to assure tight balance constraint. However the way we model imports and exports is not a fair representation of reality where decisions have to be set in advance through binding, multi-year contracts. Hence, an improvement avenue would be to introduce the notion of contract between countries or regions with e.g. predefined import/export capacities, down payments at contract set up, take-or-pay volumes and penalties throughout contract duration, and setting these as "first-stage" variables.

As we have shown in this report, the case of Belgium is relatively simple, with a small number of nodes with demand, availability of carriers or production capacity. Nevertheless, this simple case has allowed us to develop and test the model, and to identify improvement opportunities. The tool developed this far is open and offers the potential to introduce new features such as more nodes / subregions, new technologies, additional transport modes, physical constraints, ... In particular, to reflect geopolitical alliances which will have to take place above country-level in the coming years, it will be key to extend the model and cover a larger part of Europe with new / differentiated imports beyond our neighbors or the North Sea, such as pipelines from North Africa, different types of shipments from the Middle East or other regions, etc. Clearly, the approach we have undertaken in this master thesis still offers a lot of room for new work, ideas and improvements. We hope that our results and questions will inspire more research to tackle the challenges of Hydrogen Supply Chain.

Bibliography

- [1] Paolo Agnolucci and William McDowall. “Designing future hydrogen infrastructure: Insights from analysis at different spatial scales”. In: *International Journal of Hydrogen Energy* 38.13 (2013), pp. 5181–5191. ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2013.02.042>. URL: <https://www.sciencedirect.com/science/article/pii/S0360319913004084>.
- [2] A. Almansoori and N. Shah. “Design and Operation of a Future Hydrogen Supply Chain: Snapshot Model”. In: *Chemical Engineering Research and Design* 84.6 (2006), pp. 423–438. ISSN: 0263-8762. DOI: <https://doi.org/10.1205/cherd.05193>. URL: <https://www.sciencedirect.com/science/article/pii/S0263876206729185>.
- [3] A. Almansoori and N. Shah. “Design and operation of a stochastic hydrogen supply chain network under demand uncertainty”. In: *International Journal of Hydrogen Energy* 37.5 (2012). Portable Fuel Cells – Fundamental and Applications (ISPFC2010), pp. 3965–3977. ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2011.11.091>. URL: <https://www.sciencedirect.com/science/article/pii/S0360319911025894>.
- [4] Ali S. Almansoori and Nilay Shah. “Design and operation of a future hydrogen supply chain: Multi-period model”. In: *International Journal of Hydrogen Energy* 34 (2009), pp. 7883–7897.
- [5] Ahmed M. Attia, Ahmed M. Ghaithan, and Salih O. Duffuaa. “Multi-Objective optimization of the Hydrocarbon supply chain under price and demand uncertainty”. In: *International Journal of Sustainable Engineering* 14.6 (2021), pp. 1525–1537. DOI: 10.1080/19397038.2021.1995523. eprint: <https://doi.org/10.1080/19397038.2021.1995523>. URL: <https://doi.org/10.1080/19397038.2021.1995523>.
- [6] Eva H.Y. Beh et al. “Robust optimization of water infrastructure planning under deep uncertainty using metamodels”. In: *Environmental Modelling Software* 93 (2017), pp. 92–105. ISSN: 1364-8152. DOI: <https://doi.org/10.1016/j.envsoft.2017.03.013>. URL: <https://www.sciencedirect.com/science/article/pii/S1364815217303018>.

- [7] Tanveer Hossain Bhuiyan, Hugh R. Medal, and Sarah Harun. “A stochastic programming model with endogenous and exogenous uncertainty for reliable network design under random disruption”. In: *European Journal of Operational Research* 285.2 (2020), pp. 670–694. ISSN: 0377-2217. DOI: <https://doi.org/10.1016/j.ejor.2020.02.016>. URL: <https://www.sciencedirect.com/science/article/pii/S037722172030134X>.
- [8] Victor H. Cantú, Catherine Azzaro-Pantel, and Antonin Ponsich. “A Novel Matheuristic based on bi-level optimization for the multi-Objective design of hydrogen supply chains”. In: *Computers Chemical Engineering* 152 (2021), p. 107370. ISSN: 0098-1354. DOI: <https://doi.org/10.1016/j.compchemeng.2021.107370>. URL: <https://www.sciencedirect.com/science/article/pii/S0098135421001484>.
- [9] Hydrogen Import Coalition. “Shipping sun and wind to Belgium is key in climate neutral economy”. In: *Technology* 14 (2021), p. 19.
- [10] Eurostat. *Electricity price statistics*. In: (May 2022). Data extracted in April 2022.
- [11] Paolo Gabrielli et al. “Enabling low-carbon hydrogen supply chains through use of biomass and carbon capture and storage: A Swiss case study”. In: *Applied Energy* 275 (2020), p. 115245. ISSN: 0306-2619. DOI: <https://doi.org/10.1016/j.apenergy.2020.115245>. URL: <https://www.sciencedirect.com/science/article/pii/S0306261920307571>.
- [12] Grossmann I. Guillén-Goálbez G Mele F. “A bi-criterion optimization approach for the design and planning of hydrogen supply chains for vehicle use with economic and environmental concerns.” In: (2009). DOI: [dx.doi.org/10.1002/aic.12024](https://doi.org/10.1002/aic.12024). URL: <http://dx.doi.org/10.1002/aic.12024>.
- [13] Esben Hegnsholt et al. “The Real Promise of Hydrogen”. In: (July 2019). Accessed: 22-05-2022. URL: <https://www.bcg.com/publications/2019/real-promise-of-hydrogen>.
- [14] André Hugo et al. “Hydrogen infrastructure strategic planning using multi-objective optimization”. In: *International Journal of Hydrogen Energy* 30.15 (2005), pp. 1523–1534. ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2005.04.017>. URL: <https://www.sciencedirect.com/science/article/pii/S0360319905001163>.
- [15] *Hydrogen roadmap Europe : a sustainable pathway for the European energy transition and Fuel Cells and Hydrogen 2 Joint Undertaking*. Publications Office, 2019. DOI: [doi/10.2843/341510](https://doi.org/10.2843/341510).

- [16] Helgi Thor Ingason, Hjalti Pall Ingolfsson, and Pall Jensson. “Optimizing site selection for hydrogen production in Iceland”. In: *International Journal of Hydrogen Energy* 33.14 (2008). TMS07: Symposium on Materials in Clean Power Systems, pp. 3632–3643. ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2008.04.046>. URL: <https://www.sciencedirect.com/science/article/pii/S0360319908004795>.
- [17] G. Kakoulaki et al. “Green hydrogen in Europe – A regional assessment: Substituting existing production with electrolysis powered by renewables”. In: *Energy Conversion and Management* 228 (2021), p. 113649. ISSN: 0196-8904. DOI: <https://doi.org/10.1016/j.enconman.2020.113649>. URL: <https://www.sciencedirect.com/science/article/pii/S0196890420311766>.
- [18] Jiyong Kim, Younghee Lee, and Il Moon. “Optimization of a hydrogen supply chain under demand uncertainty”. English. In: *International Journal of Hydrogen Energy* 33.18 (Sept. 2008). Funding Information: This work was supported by the Ministry of Education (MOE) of Korea by its BK21 Program., pp. 4715–4729. ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2008.06.007.
- [19] Jiyong Kim and Il Moon. “Strategic design of hydrogen infrastructure considering cost and safety using multiobjective optimization”. English. In: *International Journal of Hydrogen Energy* 33.21 (Nov. 2008). Funding Information: This work was supported by the Ministry of Education (MOE) of Korea by its BK21 Program., pp. 5887–5896. ISSN: 0360-3199. DOI: 10.1016/j.ijhydene.2008.07.028.
- [20] Tianjun Liao et al. “Ant Colony Optimization for Mixed-Variable Optimization Problems”. In: *IEEE Transactions on Evolutionary Computation* 18.4 (2014), pp. 503–518. DOI: 10.1109/TEVC.2013.2281531.
- [21] Zhenhong Lin et al. “The least-cost hydrogen for Southern California”. In: *International Journal of Hydrogen Energy* 33 (June 2008), pp. 3009–3014. DOI: 10.1016/j.ijhydene.2008.01.039.
- [22] Jeff Linderoth, Alexander Shapiro, and Stephen Wright. “The Empirical Behavior of Sampling Methods for Stochastic Programming”. In: *Annals of Operations Research* 142 (Feb. 2006), pp. 215–. DOI: 10.1007/s10479-006-6169-8.
- [23] Baoding Liu. “Dependent-chance programming: A class of stochastic optimization”. In: *Computers Mathematics with Applications* 34.12 (1997), pp. 89–104. ISSN: 0898-1221. DOI: [https://doi.org/10.1016/S0898-1221\(97\)00237-X](https://doi.org/10.1016/S0898-1221(97)00237-X). URL: <https://www.sciencedirect.com/science/article/pii/S089812219700237X>.

- [24] Max Ludwig et al. “The Green Tech Opportunity in Hydrogen”. In: (Apr. 2021). Accessed: 22-05-2022. URL: <https://www.bcg.com/publications/2021/capturing-value-in-the-low-carbon-hydrogen-market>.
- [25] Matthias Niermann et al. “Liquid Organic Hydrogen Carriers and alternatives for international transport of renewable hydrogen”. In: *Renewable and Sustainable Energy Reviews* 135 (2021), p. 110171.
- [26] Joris Proost. “State-of-the art CAPEX data for water electrolyzers, and their impact on renewable hydrogen price settings”. In: *International Journal of Hydrogen Energy* 44.9 (2019), pp. 4406–4413.
- [27] Jesus Ochoa Robles, Catherine Azzaro-Pantel, and Alberto Aguilar-Lasserre. “Optimization of a hydrogen supply chain network design under demand uncertainty by multi-objective genetic algorithms”. In: *Computers Chemical Engineering* 140 (2020), p. 106853. ISSN: 0098-1354. DOI: <https://doi.org/10.1016/j.compchemeng.2020.106853>. URL: <https://www.sciencedirect.com/science/article/pii/S0098135419307161>.
- [28] Nagore Sabio et al. “Strategic planning with risk control of hydrogen supply chains for vehicle use under uncertainty in operating costs: A case study of Spain”. In: *International Journal of Hydrogen Energy* 35.13 (2010). ISMF-09, pp. 6836–6852. ISSN: 0360-3199. DOI: <https://doi.org/10.1016/j.ijhydene.2010.04.010>. URL: <https://www.sciencedirect.com/science/article/pii/S0360319910006749>.
- [29] Nikolaos V. Sahinidis. “Optimization under uncertainty: state-of-the-art and opportunities”. In: *Computers Chemical Engineering* 28.6 (2004). FOCAPO 2003 Special issue, pp. 971–983. ISSN: 0098-1354. DOI: <https://doi.org/10.1016/j.compchemeng.2003.09.017>. URL: <https://www.sciencedirect.com/science/article/pii/S0098135403002369>.
- [30] Mijndert van der Spek et al. “Perspective on the hydrogen economy as a pathway to reach net-zero CO2 emissions in Europe”. In: *Energy Environ. Sci.* 15 (3 2022), pp. 1034–1077. DOI: [10.1039/D1EE02118D](https://doi.org/10.1039/D1EE02118D). URL: <http://dx.doi.org/10.1039/D1EE02118D>.
- [31] Susara A. van den Heever and Ignacio E. Grossmann. “A strategy for the integration of production planning and reactive scheduling in the optimization of a hydrogen supply network”. In: *Computers Chemical Engineering* 27.12 (2003), pp. 1813–1839. ISSN: 0098-1354. DOI: [https://doi.org/10.1016/S0098-1354\(03\)00158-3](https://doi.org/10.1016/S0098-1354(03)00158-3). URL: <https://www.sciencedirect.com/science/article/pii/S0098135403001583>.

- [32] Young-Bin Woo and Byung Soo Kim. “A Genetic Algorithm-Based Matheuristic for Hydrogen Supply Chain Network Problem with Two Transportation Modes and Replenishment Cycles”. In: *Comput. Ind. Eng.* 127.C (Jan. 2019), pp. 981–997. ISSN: 0360-8352. DOI: 10.1016/j.cie.2018.11.027. URL: <https://doi.org/10.1016/j.cie.2018.11.027>.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl