Hydrogen infrastructure planning under uncertainty in an industrial port cluster

A robust over time approach

Meike Lafeber

Delft University of Technology

TU Delft

Hydrogen infrastructure planning under uncertainty in an industrial port cluster

A robust over time approach

by

Meike Lafeber

Student number: 4549325
Project duration: April 1, 2024 – November 18, 2024
Thesis committee:

Chair:	Prof.Dr. M.E. Warnier
First Supervisor:	Dr.Ir. P.W Heijnen
Second Supervisor:	Dr. D. Xevgenos

Cover: Unsplash (2021)

TU Delft

Preface

Over the past few months, I have dedicated myself to this thesis before you. I started this thesis because I had some skepticism about hydrogen's role in decarbonizing heavy industries, and I was motivated to challenge my own assumptions. As I delved deeper into this topic, I came to understand the pivotal role of heavy industries in our society, their deep reliance on fossil fuels and the transition that (hopefully) lies ahead.

Completing this thesis would not have been possible without the support and guidance of many people.

I would like to thank Petra, my primary supervisor, for her invaluable and straightforward advice during our weekly meetings. Your patient and personal approach helped me keep perspective during times that felt overwhelming, and your insights always left me feeling inspired and confident. I also want to thank Martijn and Dimitris for broadening my perspective and providing constructive feedback.

I am deeply grateful to my parents for their unwavering support throughout my studies, for encouraging my curiosity, and for sparking my love of complex puzzles from a young age. Additionally, thank you to all my friends for helping me step away from my work when needed and relax.

A special thank you to Ries for always being there for me and knowing exactly how to lift my spirits.

To everyone who supported me along the way, thank you! You made this journey not only more manageable but also enjoyable.

I hope you will enjoy reading this thesis.

Meike Lafeber
Delft, November 2024

Executive summary

The industrial sector is one of the most energy consuming and CO_2-emitting end-use sectors. To reach climate goals, decarbonization of these industrial sectors is imminent, however, not so apparent. Especially the hard-to-abate industry sectors, such as the chemical, mineral processing, iron, and steel sectors, are difficult to decarbonize as they require either high temperature heat or use fossil fuels as feedstock.

Hydrogen has the potential to reduce carbon emissions in industries, such as chemicals, glass, iron, and steel, as well as to serve as a cleaner heat source. To reach net-zero emissions by 2050, sectors that currently use fossil fuels for high-temperature processes and as feedstock will likely need to shift towards blue or green hydrogen. Currently, some industrial hydrogen use relies on gray hydrogen, produced from fossil fuels and contributing to emissions. In contrast, blue hydrogen captures and stores CO_2 produced from fossil sources, while green hydrogen is entirely emissions-free, generated from renewable energy. In other words, some processes need to change from gray to green/blue but most of them need to change from other fossil fuel based processes to hydrogen processes.

Hydrogen offers a way to cut carbon emissions in industries like chemicals, glass, iron, and steel, and can also act as a cleaner heat source. Achieving net-zero emissions by 2050 will likely require sectors that currently depend on fossil fuels for high-temperature applications and feedstocks to adopt blue or green hydrogen instead. Today, certain industrial applications still use gray hydrogen, derived from fossil fuels and contributing to carbon emissions. However, blue hydrogen captures and stores the CO_2 generated, while green hydrogen is emissions-free, produced using renewable energy. In essence, some processes will need to transition from gray to blue or green hydrogen, while many others will shift from fossil-fuel-based processes to hydrogen-based alternatives.

However, the timing and extent of the hydrogen transition are uncertain as they are heavily influenced by external factors such as hydrogen prices, available subsidies, and alternative decarbonization options. Additionally, industrial plant owners may be reluctant to disclose decarbonization plans due to competitive pressures, adding another layer of demand and participant uncertainty that complicates infrastructure planning.

This thesis addresses the planning of hydrogen infrastructure within an industrial port cluster (IPC). IPCs are defined by their proximity to water and concentration of industrial activities related to a specific sector. In order to effectively address spatial constraints, this thesis will plan the hydrogen networks along the current road network within IPCs. Current infrastructure planning methods have a time horizon of ten years. However, as the expectation is that the hydrogen demand will increase towards 2055, a time horizon of ten years can increase the total costs of the network when the network is implemented over time between 2025-2055. This introduces the following research question;

“How can a cost-efficient, robust pipeline network for an industrial port cluster be developed over time under uncertainty?”

To answer this question, the robust backtracking planning method (RBPM) is developed. This method aims to minimize costs over the 2025-2055 time frame while facilitating the hydrogen to the demanding plants. Because the demand for hydrogen is likely to grow over time, this method finds a robust network that is able to facilitate the demand in many possible future demand scenarios of 2055.

The robust network is then implemented incrementally for 2035, 2045, and 2055 using a backtracking approach. In this context, backtracking means that when an industrial plant transitions to hydrogen in one stage, pipelines are installed with the robust networks’ capacity, rather than just the minimum required to meet that plant’s immediate needs. This extra capacity ensures that if other plants transition in later years, the existing network can accommodate the increased demand without needing costly pipeline extensions. By preemptively building capacity, this approach reduces future installation costs and enhances the network’s ability to adapt to evolving demand patterns.

The RBPM is tested on simulations of multiple simplified IPCs. By testing different IPC simulations, it is studied how the difference in industrial plants determines the development of the network. The RBPM is compared to the results of a traditional planning approach which only plans the networks with a time horizon of ten years.

The results clearly show that the RBPM incurs lower costs over 30 years, but it requires a higher investment in 2035 due to the greater capacity installed at that time. This thesis finds that the total potential hydrogen demand and the physical size of an IPC significantly affect the performance of the RBPM compared to the traditional planning approach. Additionally, the projected installation and operational costs over time also impact the RBPM's performance relative to the traditional planning approach.

For IPCs with comparatively low hydrogen demand — typically clusters with fewer iron and steel facilities, chemical plants, or refineries — the RBPM emerges as the most economical approach. This method requires only slightly higher investment by 2035 but ultimately generates substantial savings by 2055. By installing sufficient pipeline capacity upfront, the RBPM avoids the need for additional pipelines every ten years, leading to long-term cost efficiency through 2055.

For IPCs with high hydrogen demand — typically found in iron and steel plants, basic chemical plants, or refineries — the initial installation costs and ongoing operational expenses of RBPM make it less advantageous. While RBPM may offer slightly better economic profitability over a 30-year period, the substantial investment required in 2035 compared to traditional planning makes implementation challenging due to budget constraints. In these high-demand clusters, the decision between RBPM and the traditional approach for developing a hydrogen pipeline network depends on the cluster's budget, anticipated future installation costs, and projected operational expenses over time.

Opportunities for further research include the application of the RBPM to a real case study to validate the result, increasing the amount of possible future scenarios by incorporating uncertainty in installation and operating costs and increasing the demand and participant uncertainty range. Lastly, another research direction to explore is the generation of different robust network methods and their performance.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

Abbreviation	Definition
ABM	Agent-Based Modeling
CCUS	Carbon Capture Usage and Storage
IDPM	Immediate Demand Planning Method
IPC	Industrial Port Cluster
MCStT	Minimum Capacitated Steiner Tree
MCST	Minimum Capacitated Spanning Tree
ML-RBPM	Minimum Length Robust Backtracking Planning Method
MST	Minimum Spanning Tree
MStTG	Minimum Steiner Tree in a Graph
MILP	Mixed-Integer Linear Programming
RBPM	Robust Backtracking Planning Method
SEC	Specific Energy Coefficient
TRL	Technology Readiness Level
ONLT	Optimal Network Layout Tool

Symbols

Symbol	Definition	Unit
$C(t)$	Total costs in euros for one timestep t, combining capital and operational costs	[M€]
$C^C(t)$	Capital costs in euros for one timestep t	[M€]
$C^O(t)$	Operational costs in euros for one timestep t	[M€]
$C^m_{l,c,s}(t)$	Costs per decade for layout l, configuration c, and scenario s using method m	[M€]
$CC^m_{l,c,s}(t)$	Cumulative costs by the decade t for method m, layout l, configuration c, and scenario s	[M€]
c_c	Cluster coefficient	-
d_{ij}	Diameter of one pipeline	[mm]
$D_{n,t}$	Hydrogen demand for each industrial plant n at timestep t	[ton H_2/day]
$D_{\max,n}$	Maximum hydrogen demand as a function of production capacity	[ton H_2/day]
$D_{s,t,n}$	Demand for each scenario s, timestep t, and terminal node $n \in N_{\text{terminal}}$	[ton H_2/day]
$D_{\text{est},n}$	Estimated hydrogen demand if the plant n transitions to hydrogen	[ton H_2/day]
Δ_D	Demand range, defines deviation from $D_{\max,n}$	[ton H_2/day]

Continued on next page

Table – continued from previous page

Symbol	Definition	Unit
E	Set of edges of a graph	-
f_{OC}	Fraction of operational costs, used in cost sensitivity analysis	-
$FD_{G_{i,t},s}$	Facilitated demand for network G_i at timestep t	[ton H_2/day]
$FD^m_{l,c,s}(t)$	Facilitated demand for network $N^m_{l,c,s}$ at timestep t	[ton H_2/day]
G_t	Hydrogen network at timestep t	-
$G_t(N, E)$	Capacitated Steiner tree for timesteps t, where $G_t(N, E) \subseteq H(N, E)$	-
$H(N, E)$	Road network depicted as a Steiner graph	-
$L^m_{l,c,s}(t)$	Total network length under method m for layout l, configuration c, and scenario s at timestep t	[m]
l	Identification of layout number	-
l_{ij}	Length of one edge	[m]
n	Industrial node, an element of N_{terminal}	-
N	Nodes in a graph or network	-
N_{steiner}	Steiner nodes, a subset of all the nodes in a Steiner graph	-
N_{terminal}	Terminal nodes, a subset of all the nodes in a Steiner graph	-
$NFD^m_{l,c,s}(t)$	Non-facilitated demand for network $N^m_{l,c,s}$ at timestep t	[ton H_2/day]
p_{extra}	Probability of an additional industrial plant appearing	-
p_{highT}	Probability of a plant transitioning at high temperature or with a feedstock process	-
p_{lowT}	Probability of a plant transitioning at low temperature	-
p_{midT}	Probability of a plant transitioning at medium temperature	-
p_T	Transition uncertainty	-
$p_{n,t}$	Transition probability of an industrial node n at timestep t	-
$ProdCap_n$	Production capacity of a process	[ton/year]
q_{ij}	Capacity of one pipeline	[ton H_2/day]
$RC^m_{l,c,s}(t)$	Relative costs between method m and RBPM by timestep t	[M€]
s	Scenario number	-
$SEC_{technology,n}$	Specific Energy Consumption of the technology	[MJ/kg H_2]
t	Timestep indicating the end of one planning decade	-
TRL_n	Technology readiness level for an industrial node n	-
$\Phi_{G_{i,t},D_{s,t}}$	Maximum flow that network i can facilitate for demand scenario $D_{s,t}$	[ton H_2/day]

1 Introduction

1.1. Problem Statement

The global average temperature peaked in 2023, at 1.45 degrees above pre-industrial levels (Organization, 2024). This is an indication of a global trend: the world is warming. Global warming has far-reaching effects, which include a decline in biodiversity, a rise in natural disasters and extreme weather occurrences, and an increase in related societal risks (IPCC, 2022). The emission of greenhouse gases, particularly CO_2, contributes to global warming. The United Nations established the Paris Agreement in 2015 to limit global warming, aiming for net-zero emissions by 2050 (United Nations, 2015). Additionally, the European Union has elevated its objective to decrease GHG emissions by at least 55% compared to the 1990 level by 2030 as part of the Fit-For-55 package, in pursuit of its 2050 target (European Council, 2024).
The industrial sector is one of the most energy-consuming and CO_2-emitting end-use sectors (IEA, 2023b). In 2022, the Dutch industrial sector emitted 30% of the total annual CO_2 emissions in the country (Rijksoverheid: Emissieregistratie, 2024). Therefore, in order to reach the EU targets, the industrial sector needs to decarbonize as well. The decarbonization of the industrial sector is therefore, very important but no so apparent. Especially the hard-to-abate industry sectors, such as the chemical, mineral processing, iron, and steel sectors, are difficult to decarbonize as they require either high temperature heat or use fossil fuels as feedstock.

1.1.1. Hydrogen as decarbonization technology in industry

Hydrogen has the potential to reduce carbon emissions in hard-to-abate industries, such as the chemical, glass, iron, and steel sectors, and as a substitute for heat generation (Deloitte, 2023; Namazifard et al., 2024; Vine, 2021). The subsectors with the potential for transitioning to hydrogen either utilize hydrogen as a feedstock for decarbonization (e.g., basic chemicals, iron, and steel) or employ hydrogen as a heat source.

Hydrogen plays a pivotal rola as a feedstock in multiple industrial subsectors, particularly within the chemical and iron and steel industries. Hydrogen is crucial for the synthesis of ammonia, methanol, and olefins in the basic chemicals sector. Ammonia manufacturing often employs the Haber-Bosch process, which synthesizes hydrogen and nitrogen at elevated pressure. Currently, gray hydrogen, derived from natural gas, predominates this process (Cioli et al., 2021). Nonetheless, a shift to green hydrogen, produced through electrolysis utilizing renewable energy, or blue hydrogen, generated from natural gas with Carbon Capture and Storage (CCUS), provides a near-zero carbon footprint for ammonia production (Rouwenhorst et al., 2021). In this thesis, the term ”transition of hydrogen” refers to the shift towards clean hydrogen, namely blue or green hydrogen.

In the iron and steel production, the primary method of producing crude steel rely heavily on coal and coke, generating substantial CO_2 emissions. The use of hydrogen in direct reduction processes, where iron ore is converted to iron using green hydrogen instead of coal, can potentially reduce emissions by up to 98% (IEA, 2024a). In 2024, the technology has been tested on multiple full scale prototypes (IEA, 2024a). This hydrogen-based technology, while promising, is not yet prepared for commercial implementation, and the timeline for its readiness for full-scale deployment remains uncertain (Fan & Friedmann, 2021).

In industrial operations, heat is necessary to initiate chemical reactions or to accomplish certain physical and chemical changes. The type of heat required for typical operations varies from 50 C to 2000 C (Neuwirth et al., 2022b). Conventional heat-generating techniques frequently entail burning fossil fuels, which releases greenhouse gases into the atmosphere. The temperature of the heat, dictates what kind of decarbonization technologies are available to replace the fossil fuel heat generation technique; the higher the temperature needed for the process, the less clean heat solutions are available. For the high temperature heat processes only hydrogen, biomass, and Carbon Capture Usage and Storage (CCUS) are available (Vine, 2021).

How fast and what kind of decarbonization technologies will be implemented by industrial plants, depends on multiple factors. These factors include political landscape, economical factors, social factors, environmental factors, technologies available, and available infrastructure (IEA, 2023b). As it depends on all these factors, the specific decarbonization pathway for an industrial plant is not easily predicted. However, if a plant has a desire to transition to hydrogen, sufficient hydrogen infrastructure is a prerequisite.

1.1.2. Hydrogen transmission infrastructure

In order to facilitate a hydrogen transition, infrastructure is needed to distribute hydrogen from the production or import site to the industrial plant. The network of preference to satisfy industrial demand is a pipeline transmission network, as it outperforms other alternatives like transportation as volume increases and can easily provide a constant flow of hydrogen for industrial processes (André et al., 2013). A hydrogen pipeline network is not yet widely implemented, but, by the need for rapid decarbonization, networks are now widely pushed to be implemented. Historically, gas pipeline networks evolved incrementally, growing node-by-node with increasing demand, leading to sub-optimal designs . Hydrogen networks can now be designed with the advantage of leveraging advanced computational methods to optimize the design and test configurations (Hammond et al., 2024). A hydrogen network is often visualized by supply nodes (sources) and demand nodes (sinks) who are connected with pipelines. These pipelines must allow the transportation of hydrogen from sources to sinks; the volume of flow a pipeline can manage is referred to as its capacity.

Deep uncertainty

One of the challenges in constructing a hydrogen network is in the deep uncertainty surrounding future hydrogen demand and supply, which is contingent upon technical, political, and economic factors influencing the transition of fossil fuel-based technologies (IEA, 2023a). Deep uncertainty means that because of the complexity of the system, there are a large number of plausible futures that cannot be ranked based on the likelihood of their occurrence (J. H. Kwakkel et al., 2010). Although, it is difficult to predict the future hydrogen demand for industrial plants, it is imperative to consider the future when planning the pipeline diameter and location. Since pipelines typically have a lifespan between 30-50 years and an implementation time of several years, the future hydrogen demand should be taken into account when planning hydrogen networks, in order to have a cost-efficient network in 2050 (Khan et al., 2021).

1.1.3. Industrial Port Clusters

This thesis will concentrate on the development of hydrogen networks within industrial port clusters. Specifically, a focus is set on clusters within particular industrial subsectors where hydrogen presents a feasible solution for decarbonization.

An industrial port cluster (IPC) refers to an industrial cluster that is geographically located near a port. An industrial cluster consists of a group of geographically adjacent, interconnected companies and related institutions in a specific field (Kim et al., 2023). Clusters are characterized by specialization in related industries and the dynamic interactions among many firms, rather than the presence of a single large firm or plant or the specialization on only one narrow activity (European Observatory for Clusters and Industrial Change, 2020). The advantage of having an industrial cluster near a port is the convenient access to resources and suppliers, as the port can handle a significant amount of import and export (Pivetta et al., 2024). However, the presence of water and the high concentration of firms result in limited space for infrastructure expansion. This poses a challenge for infrastructure planners, as they have to consider physical boundaries and constraints in the form of waterways and private property.

Actors

Within an IPC hydrogen network, the following primary actors are recognized:

- The network operator, who is tasked with network planning and bears the costs associated with its development and maintenance (GasUnie, 2024). The network operator is frequently semi-private. The network operator's interests involve ensuring user facilitation and a safe transition while generating profit.
- Industrial plant owners, who are represented by a plant owners association. They each determine the timing and conditions under which the plant will alter its process technology for decarbonization. The plant owner is motivated by anticipated profit. The owner lacks a direct incentive for a decarbonized society but can be influenced by economic policies such as carbon taxes, subsidies, or energy carrier prices. The plant owners' association advocates for adequate IPC infrastructure to meet future plant requirements while minimizing utilization costs for current and prospective customers (Deltalinqs, 2024).
- Hydrogen producers, who function as either hydrogen producers or importers. Producers need a solid infrastructure and reliable demand linked to it to maintain a viable business model and achieve economic benefits (IEA, 2023a).
- The Port Authority, who is responsible for managing, operating, and developing the port area. The Port Authority, despite being publicly owned, functions with a profit-driven approach. It aims to meet European Union standards and seeks to become future proof port, which means to be environmentally sustainable (Port of Rotterdam, 2024).

The importance of cost-efficiency is significant for all stakeholders, whether it is to reduce investment costs, or to have low utilization prices. This creates the need for a cost-effective hydrogen network. It is also important for the hydrogen network to effectively facilitate both current and future users. Finding a balance between these two requirements is difficult, as the most cost-effective network is having no network at all, but that would not satisfy the needs of any users and hinder the shift towards a near-zero emissions cluster. This highlights the need for a decision-making tool that helps in balancing cost and demand facilitation in developing a hydrogen network in an IPC.

1.2. Identification of the research gap

For identification of the knowledge gap a literature study is conducted, where sixteen papers have been found. The development and optimal design of hydrogen networks is a state-of-the-art research field. As can be seen in table 1.1, five articles were recently (2024) published about hydrogen pipeline network planning. Fourteen out of sixteen articles are analyzing the European hydrogen network; which could be caused by the outspoken ambition of Europe to have an independent, reliable clean energy system (Europese Commissie, 2022). Below, we will inspect the scale of the studies, how they deal with robustness, and how they investigate the development of a network over time.

Scale

Most of these studies simulate hydrogen networks at the national or regional scale, where one node represent larger geographical areas such as provinces, hubs, or industrial clusters. In these models, the edges represent pipeline or truck connections between these nodes. However, there is a notable scarcity of studies focusing on the local scale, where each node represents a single site. Furthermore, in the papers focusing on regional and national scale, the pipeline locations are not taken into account as a decision variable. In every article, the pipeline location is predefined in advance, for example, by calculating the Euclidean distance between nodes and multiplying this length with a detour factor(Tlili et al., 2020). Moreover, in some articles, in order to reduce the solving time, only subsets of possible pipeline connections between nodes are incorporated. By predefining the distance and neglecting the precise location of pipelines, the flexibility of determining the location of pipelines to minimize costs is not exploited by these papers, which affect the cost efficiency of these networks (Bolat & Thiel, 2014).

Two studies are found that researched local hydrogen networks. André et al. (2014) developed a heuristic to derive a minimum cost hydrogen pipeline network for fueling stations, and Hammond et al. (2024) derived a minimum cost hydrogen pipeline network considering obstacles. In these local scale studies, one node represents one single site and the pipeline lengths are the actual real lengths.

Table 1.1: Overview literature on hydrogen networks

Source	Scope	Scenarios	Area	Over-time aspect
(Beagle et al., 2024)	regional	7000	Texas	
(Efthymiadou et al., 2024)	regional	1	Great Britain	x
(Hanto et al., 2024)	international	3	Europe	x
(Namazifard et al., 2024)	regional	3	Belgium	x
(Erdoğan & Güler, 2023)	national	5	Turkey	x
(Caglayan et al., 2021)	national	1140	Europe	
(Husarek et al., 2021)	national	8	Germany	
(Ochoa Robles et al., 2019)	regional	6	France	x
(Reuß et al., 2019)	regional	Many	Germany	
(Welder et al., 2018)	regional	3	Germany	
(Yáñez et al., 2018)	regional	2	Spain	x
(Moreno-Benito et al., 2017)	regional	7	Great Britain	x
(André et al., 2013)	local, regional	1	France	
(Almansoori & Shah, 2011)	regional	9	Great Britain	x
(Hammond et al., 2024)	local	1	Great Britain	
(André et al., 2014)	regional	2	France	x

However, these studies do not take into account the deep uncertain nature of hydrogen demand.

Deep uncertainty

Based on the found articles, the most influential and uncertain parameters of the network are found to be the hydrogen demand, supply and price. To deal with this demand or supply uncertainty, most articles use the best-guess or most likely approach, where only a few price or cost scenario's are explored in order to predict future systems. As explained in Section 1.1.2, a few scenarios cannot capture the future of the hydrogen system due to the deep uncertain nature of the system (J. H. Kwakkel et al., 2010).

Three studies ran many scenario's (>100) either in demand or supply to grasp the behavior of the system, deal with the deep uncertainty, and create a robust analysis (Beagle et al., 2024; Caglayan et al., 2021; Reuß et al., 2019). For example, Beagle et al. (2024) studied the effect of different policies on the Texas hydrogen system and ran 1000 Monte Carlo simulations per policy to address the price uncertainties. These studies show that deep uncertainty can be addressed in hydrogen network planning with the use of many scenario runs, and thereby creating a robust solution that is insensitive to a range of plausible futures (Maier et al., 2016).

Over time development

The three studies that addressed high uncertainty by examining numerous scenarios predicted a robust system for 2050. Nonetheless, these studies did not consider the implementation of this robust system over time. André et al. (2014) studied the over time implementation of a hydrogen network leading to 2050, highlighting the necessity for more temporal deployment studies in the hydrogen infrastructure literature.

Research gap

The knowledge gap that exist in the literature is the lack of studies on local, robust hydrogen network planning that look at the development of the hydrogen network over time. All three aspects are considered separately in studies, however, a combination of the aspects deep uncertainty, local scale and temporal deployment has not been studied for hydrogen networks.

1.3. Research question

The identified research gap within the literature is the lack of studies regarding a robust planning approaches that consider over time development of the network under deep uncertainty. This thesis

focuses on the development of a hydrogen network in an industrial port cluster (IPC), emphasizing the uncertainties related to demand and participant involvement to create a robust network over time. This thesis aims to answer the research question;

> **How to plan over time a cost-efficient, robust pipeline network for an industrial port cluster under uncertainty?**

In this thesis, a method is developed to plan a hydrogen network over time, taking into account cost efficiency and robustness. This method is specifically designed for IPCs; however, it is not applied to a single case study IPC. Instead, a general IPC model is developed to test the proposed method under various conditions. Before addressing the main research question, the following sub-questions are explored.

1. What characteristics of Industrial Port Clusters should be considered for the hydrogen network development?
2. How can cost-efficiency and robustness be operationalized in a physical network?
3. What is a suitable method to generate a robust and cost-efficient hydrogen network over time?
4. How does the robust over time planning method perform compared to simpler planning methods?
5. What is the performance of the developed method across different clusters?
6. How does the developed method perform under different demand uncertainty ranges?

1.3.1. Thesis outline

In chapter 2, the first sub-question is investigated by defining characteristics of an IPC. Then, chapter 3 aims to answer the second sub-question by creating a theoretic framework regarding network optimization and deep uncertainty, and at the end of this chapter, the concepts of cost-efficiency and robustness are operationalized. Next, chapter 4 answers sub-question three by describing the developed method. In chapter 5, the data is presented that is needed to analyse the developed method, and in chapter 6 the presented method is analyzed where sub-question 4, 5, and 6 are answered by conducting three separate experiments. The findings of the results and limitations are further discussed in chapter 7, and conclusions and future research are presented in chapter 8.

CoSEM master thesis

This thesis is written to obtain the MSc degree for Complex Systems Engineering and Management (CoSEM), as this thesis addresses a design problem in a complex system with multiple actors and uncertainties. First of all, the situation of a future hydrogen network in the port is highly complex due to the different technical and economic factors playing a role. Second, the system involves numerous actors, including ports, network operators, and industrial plants, making it a multi-actor problem involving both public and private interests. Because of this complexity, there is not one solution for this problem (Johannesson & Perjons, 2021).

2

Important aspects of an IPC

This thesis specifically focus on the hydrogen network development in an industrial port cluster (IPC). However, this research does not assess only one real-life IPC example, but examines the network development on a multitude of modeled IPCs to develop a generic method. Therefore, in order to build a hydrogen network development method, it is important to explore the characteristics of an IPC and how they relate to the hydrogen network. In this chapter, firstly, the IPC is defined further, secondly, the future role of hydrogen in an IPC is explored. Lastly, the important aspects of hydrogen network development in an IPC are explained including an actor analysis to derive the network objectives.

2.1. What is an IPC?

The term Industrial Port Cluster (IPC) is defined for this thesis as an industrial cluster geographically close to a port. First the term industrial cluster is explored, after which the port and specifically its physical characteristics are described. Both the physical properties of a port and the characteristics of an industrial cluster are important to consider for the planning of a hydrogen network.

2.1.1. Industrial cluster

There are multiple definitions of clusters and industrial clusters in the academic literature, which mainly differ in their emphasis on cooperation or geographical scope (Boja, 2011). For instance, Porter (2000) defines clusters as *"geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions, in a particular field that compete but also cooperate"*. This definition highlights the dual nature of clusters, where firms not only compete but also collaborate within a shared environment. However, the geographic scope relates to the distance over which informational, transactional, incentive, and other efficiencies occur.

For this thesis, the industrial cluster definition of Morosini (2004) is used, who characterizes an industrial cluster *"as a community of people and a population of firms with economically linked activities localized in close proximity in a specific geographic region, where a significant part of the population cooperates"*. Morosini's definition emphasizes the importance of geographical closeness, highlighting that in industrial clusters firms also benefit from reduced transportation costs by having the firms up and down the supply chain be close together or by managing one feedstock together. Additionally, in industrial clusters firms frequently work together to (re-)use residual flows, such as water, energy, and materials.

Besides residual flow sharing, firms within an industrial cluster also engage in other types of collaboration. Firms collaborate on linked activities, sharing and nurturing a common stock of products, technology, and organizational knowledge to generate superior products and services in the marketplace (Mortensen et al., 2023). However, there is also competition between these firms. As a result, the firms do not want to share all information with each other. As firms do not want to disclose their possible future transition pathways, it is difficult to plan ahead for hydrogen network developers.

2.1.2. Physical port aspects

The advantage of having an industrial cluster near a port is the convenient access to resources and suppliers, as the port can handle a significant amount of import and export (Pivetta et al., 2024). The term port typically refers to a network or infrastructure that enables the loading and unloading of cargo from vessels and facilitates the transfer of cargo between different transportation modes. A port has a water infrastructure with at least the following essential facilities according to Koningsveld et al. (2023): approach channel(s), maneuvering areas, approach channels, and berths.

The water infrastructure is of influence for the development of a hydrogen pipeline network, as pipelines placement under water is more difficult and therefore more expensive. However, it is outside the scope of this thesis to fully study the workings of water infrastructure in a port. Therefore, an assessment of the wet infrastructure in major European industrial ports was conducted to look for patterns. The visual assessment of European industrial port clusters involved examining Google Maps satellite pictures. These pictures are in small presented below. For a more zoomed in view of the IPC overview, see appendix B.

Wet infrastructure

A visual inspection of ports such as Rotterdam, Duisburg, Antwerp, and Dusseldorf shows the water infrastructure in blue. These four ports are chosen since these ports are the biggest industrial ports in the European Union (Deloitte, 2023). First, of all one can see that the water infrastructure looks like a tree that does not have any cycles. Having a cycle would also be inconvenient as it would create an island and thereby decrease the accessibility of the facilities located on this island.

Zooming in on part of the ports, it can be seen that the wet infrastructure typically consists of multiple docks where ships can load and unload, which corresponds with the findings of Koningsveld et al. (2023). The extent and layout of these docks vary depending on the port. This is related with the kind of ships that should be able to dock and thus turn in the port (Koningsveld et al., 2023). Additionally, turning basins, used for ship maneuvering, are often located at the entrance of these docks.

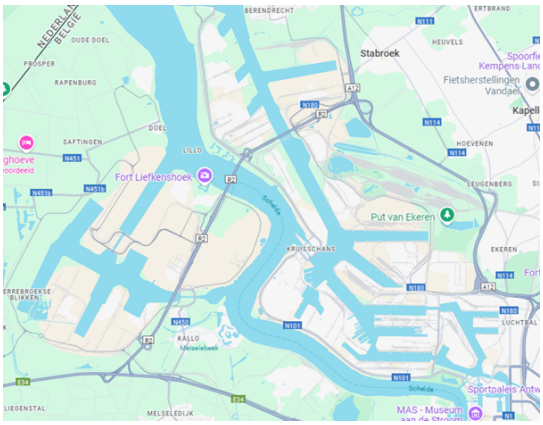
Road network

Looking at other infrastructure in the ports, one can see train rails and roads. All this infrastructure is needed to distributed the loaded or produced goods to the hinterland. Focusing on the road network, one can see that the roads barely cross the water, per port there are one or two bridges crossing the water. Additionally, the area between docks is often accessed by one road with occasional branches. The primary purpose of this road network is to provide access to the plants and facilitate the efficient transport of produce to other locations. Consequently, the roads are designed to be direct, minimizing unnecessary turns that could increase transit time.

Furthermore, the presence of water and the high concentration of firms in the IPC result in limited space for infrastructure expansion. This poses a challenge for hydrogen network development, as physical boundaries and constraints in the form of waterways and private property will need to be considered when planning new pipelines. In the Port of Rotterdam, the pipeline infrastructure has a dedicated channel, that is placed next to the national road (CES Rotterdam-Moerdijk, 2021). Therefore, in this thesis, it is assumed that the pipelines will be placed next to the roads. Consequently, the pipelines will not be hindered by private property. Additionally, it is assumed that the pipelines will not be placed next to a bridge,

(a) Port of Rotterdam

(b) Port of Duisburg



(c) Port of Antwerp

(d) Port of Dusseldorf

Figure 2.1: Visual inspection of different ports made with Google (2024)

2.2. Hydrogen in an IPC

Hydrogen has the potential to reduce carbon emissions in hard-to-abate industries, such as the chemical, iron, and steel sectors, and as a substitute for heat generation (Deloitte, 2023; Namazifard et al., 2024; Vine, 2021). The subsectors with processes that can transition to hydrogen include those that can achieve decarbonization by using hydrogen either as a feedstock (e.g., basic chemicals, iron, and steel) or as a heat source.

Hydrogen as feedstock

Hydrogen plays a pivotal role as a feedstock in various industrial subsectors, especially in the chemical and iron and steel industries. In the basic chemicals sector, hydrogen is essential for synthesizing ammonia, methanol, and olefins. For example, ammonia production typically involves the Haber-Bosch process, which combines hydrogen with nitrogen under high pressure. Currently, gray hydrogen, produced from natural gas, dominates this process. However, a transition to green hydrogen, derived from electrolysis powered by renewable energy, or by blue hydrogen, derived from natural gas but with Carbon Capture and Storage (CCUS) offers a near-zero carbon footprint for ammonia production (Rouwenhorst et al., 2021). Similarly, methanol production, a basic chemical, which uses hydrogen in the Auto Thermal Reforming process, can benefit from green hydrogen to produce CO_2-neutral methanol. This shift is crucial for achieving climate goals in this sector (Cioli et al., 2021).

In the iron and steel production, hydrogen's role as a feedstock can be transformative. Traditional methods rely heavily on coal and coke, generating substantial CO_2 emissions. The use of hydrogen in direct reduction processes, where iron ore is converted to iron using electrolytic hydrogen instead of coal, can potentially reduce emissions up to 98% (IEA, 2024b). Despite the high costs and limited current implementation, this technology represents a significant advancement towards decarbonizing steel production. By integrating green hydrogen into these processes, industries can make substantial progress towards sustainability and meet regulatory and environmental targets more effectively.

Hydrogen to generate heat

In industrial operations, heat is necessary to initiate chemical reactions or achieve certain physical and chemical changes. The type of heat required for each operation varies from $50\,^{\circ}\mathrm{C}$ to $1600\,^{\circ}\mathrm{C}$

(Neuwirth et al., 2022b). Conventional heat-generating techniques often involve burning fossil fuels, which releases greenhouse gases into the atmosphere.

The temperature of the heat dictates which decarbonization technologies can replace fossil fuel heat generation. For low-temperature ($\leq 150\,^{\circ}\mathrm{C}$) and mid-temperature heat ($150\,^{\circ}\mathrm{C}$ - $500\,^{\circ}\mathrm{C}$), there are several clean heat solutions. Direct electrification is one of the most promising options if the electricity grid can handle it and electricity prices can compete with fossil fuel prices (IEA, 2024a; Wei et al., 2019). However, for high-temperature processes ($\geq 500\,^{\circ}\mathrm{C}$), there are fewer options (see Table 2.1). Clean heat solutions for high-temperature applications can be generated from biodiesel or using CCUS, thus limiting carbon emissions. However, both face challenges: biomass production competes with food production for land, and CCUS requires new infrastructure; currently, the legal framework in the ETS does not accommodate CCUS (Vine, 2021; Wei et al., 2019). Another high-temperature option is heat generated by hydrogen technology, which can reach $2800\,^{\circ}\mathrm{C}$, but, like CO_2, it requires new infrastructure (Vine, 2021).

Table 2.1: Possible clean heat solutions and to which process they can be applied [simplified table adopted from Vine (2021)]

	Solar thermal	Geothermal	Nuclear	Electrification	Hydrogen	Biomass and biofuels	CCUS
Low temperature heat process	x	x	x	x	x	x	x
Mid temperature heat process		x	x	x	x	x	x
High temperature heat process					x	x	x

2.2.1. Industrial Subsectors Suitable for Hydrogen-Based Decarbonization

As defined before, an industrial cluster is already defined as geographically close firms with linked economic activity. For this thesis it is assumed that plants who belong to the same industrial subsector have linked economic activity. In this section we will scope the industrial subsectors that will be considered this thesis. The European Observatory for Clusters and Industrial Change (2020) identifies 51 specialized subsectors for industrial sectors in their performance study on European regional industrial clusters. This thesis will focus on eight industrial subsectors identified as likely candidates to transition to hydrogen as decarbonization strategy (Neuwirth et al., 2022b). These subsectors are;

- Basic chemicals
- Refineries
- Iron and Steel
- Metal processing
- Paper and pulp
- Glass
- Mineral processing
- Non-ferrous metals

These subsectors are considered because their production processes include key steps that can be transitioned to hydrogen-based technologies as part of their production operations (Neuwirth et al., 2022a). These processes can either be transitioned to a hydrogen clean heat solution, or with hydrogen as feedstock.

2.2.2. Process specifics

In this thesis, we argue that the probability for hydrogen transition can be determined by the specification of the individual processes and technology within each plant, rather than the effect of the whole complex system on the plant. In Figure 2.2, a schematic overview illustrates the relationship between subsectors, processes, and process specifics of a plant. A plant belongs to a subsector and has a process that can be decarbonized by adapting the hydrogen decarbonization technology. Additionally, the process temperature and the technology readiness level (TRL) are indicated, and their relation to the concepts. Below, it is discussed how process temperature and TRL are related to the probability of a process adapting to the decarbonization technology.

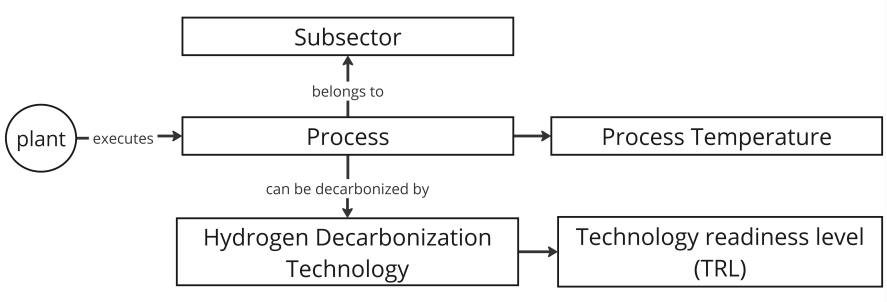


Figure 2.2: Schematic representation depicting the connections between subsectors, processes, and decarbonization technologies. This diagram categorizes each plant within its respective subsector and specifies the processes that can transition to hydrogen-based technologies. Furthermore, it mentions the process temperature and technology readiness level (TRL) that influence the adaptation of the decarbonization technology.

Process temperature

The process temperature indicates the required temperature needed for the process. As explained in section 2.2, clean heat solutions are available. Assuming that by 2050 all process heat is generated by clean heat solutions, the question arises: which clean heat solution will be implemented? In this thesis, the likelihood of adopting hydrogen decarbonization technology is linked to the number of competing heat technologies that hydrogen faces. High temperature heat is expected to have the highest probability for adaptation, as it only has two competing technologies. The assumption that a portion of high temperature heat processes is likely to transition to hydrogen is supported by various studies (IEA, 2021; Namazifard et al., 2024; Vine, 2021; Wei et al., 2019).

What is left out of consideration are the external factors such as economical, political or social factors on which clean heat technology can be adapted. Such factors can give one clean heat technology a head-start or a boost over the others. In this thesis, these factors and their influence on hydrogen technology adaptation is later in this thesis incorporated by introducing uncertainty in the transition probability of a process.

Technology readiness level

The technology readiness level (TRL) represents the maturity level of the technology. The TRL is an ordinal scale ranking from level 1 to level 9. TRL 1 represents the initial idea of a technology, TRL 4-6 represent the prototype phase, and TRL 9 indicate that the technology is commercially available. In Table 2.2, all TRL levels are described (IEA, 2019).

The TRL of a hydrogen decarbonization technology also influences the likelihood of transition in the future. A technology that is currently in the early stages of prototyping is unlikely to be commercially available and put into use within ten years. Therefore, for this thesis, the TRL is seen as an indicator for the implementation of a technology in the near future (ten years). Beyond this period, the commercial maturity of the technology remains uncertain due to various influencing factors. Therefore, the TRL assigned now (2024) is not considered to affect the likelihood of transition after ten years.

Table 2.2: Representing the Technology Readiness Levels and their description adapted from IEA (2019)

Level	Level name	Description
1	Initial idea	Basic principles have been defined
2	Application formulated	Concept and application of the solution have been formulated
3	Concept needs validation	Solution needs to be prototyped and applied
4	Early prototype	Prototype proven in test conditions
5	Large prototype	Components proven in conditions to be deployed
6	Full prototype at scale	Prototype proven at scale in conditions to be deployed
7	Pre-commercial demonstration	Solution working in expected conditions
8	First-of-a-kind commercial	Commercial demonstration, full-scale deployment in final form
9	Commercial operation in relevant environment	Solution is commercially available, needs evolutionary improvement to stay competitive

2.3. Hydrogen network development in an IPC

As processes transition to hydrogen-based technologies, sufficient hydrogen supply and infrastructure are essential to meet this growing demand. In this section, the hydrogen supply chain is discussed. Following this, the current state of hydrogen network planning is discussed.

2.3.1. Hydrogen supply chain

The hydrogen supply chain encompasses the production, transportation, storage, and utilization of hydrogen, as illustrated in Figure 2.3 (Griffiths et al., 2021). For this thesis, the focus is on local distribution within an IPC. The hydrogen network within an IPC consists of interconnected supply, demand, and storage locations. Below these four components shall be shortly explained, as well as important assumptions connected to these components.

Hydrogen supply

Hydrogen supply in an IPC can originate from several sources, including generation plants, electrolysis, steam methane reforming, and import facilities. Import facilities may involve pipelines or ammonia cracking plants, which convert incoming ammonia from tankers into hydrogen (Griffiths et al., 2021). The primary objective of the hydrogen network is to efficiently distribute hydrogen from these supply nodes to the industrial plants that consume it.

Hydrogen storage

In addition to supply, storage is another crucial element of the hydrogen network in an IPC, acting as a buffer to ensure a consistent supply of hydrogen, particularly when hydrogen is imported via tankers. Effective hydrogen storage is vital due to the gas's low density (Parolin et al., 2022). Various storage methods can be employed, including compressed gaseous hydrogen and liquid hydrogen, each suited for different applications. Additionally, underground storage solutions provide efficient options for larger capacities, enabling a reliable and consistent supply of hydrogen within the network.

In this thesis, the focus is on the facilitation of the demand by the hydrogen network. Hereby it is assumed that there is enough hydrogen supply in the system to provide this. Furthermore, storage is not considered as a separate location. It is assumed that the storage facilities are located at the same location as supply or demand facilities. Consequently, for this thesis the scope is on planning a hydrogen network in such a way that the pipelines can facilitate the demand.

Hydrogen distribution infrastructure

This thesis primarily addresses the distribution of hydrogen. Gaseous hydrogen can be transported over land through either pipelines or trucks. Pipelines are particularly well-suited for the delivery of substantial quantities of hydrogen; they represent the most economically viable option for distributing more than 1 ton of H_2 per day over distances up to 3000 km (Korner, 2015; Parolin et al., 2022). In contrast, when the quantity of hydrogen distributed is less than 1 ton, trucks are the more feasible alternative. Given that this thesis focuses on industrial clusters that are likely to require large and continuous flows of hydrogen, pipelines emerge as the most appropriate method of distribution.

Hydrogen demand

The future outlook for hydrogen demand has been addressed earlier in this thesis, highlighting the potential for various processes to transition to hydrogen-based solutions, see section 2.2.1. The process specifics play a critical role in assessing the likelihood of this transition in the future.

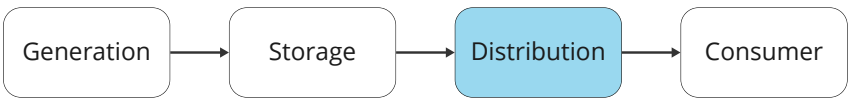


Figure 2.3: Schematic overview Hydrogen Supply chain, and the place of distribution in this supply chain

2.4. Actor analysis

In order to design an artifact that can help plan a hydrogen network, it is important to look at the different actors and stakeholders involved in planning such network. The important actors are the network operator, the industrial plant owners, the hydrogen producers, and the port authority who have been introduced in Section 1.1.3. An overview of the actors, their description and their interests can be seen in Table 2.3. This section covers the requirements of the hydrogen network based on the actors' interests.

Table 2.3: Overview of the direct stakeholders and their interests regarding a hydrogen network in an IPC.

Actor	Description	Interests
Network operator	Semi-private entity responsible for network planning, development, and maintenance. Facilitates users and ensures a safe transition.	- Economic profit - Safety of network - Facilitating users - Facilitating NZE transition
Industrial plant owners	Represented by a plant owners association; determine plant changes for decarbonization. Influenced by policies such as carbon taxes.	- Economic profit - Sufficient infrastructure
Hydrogen producers	Function as producers or importers; require solid infrastructure and reliable demand to maintain a viable business model.	- Economic profit - Hydrogen market reachable by infrastructure
Port Authority	Publicly owned entity managing, operating, and developing the port area. Aims to meet EU standards and become a future-proof, sustainable port.	- Decrease emissions - Adhere to EU regulations - Safe environment - Keep industrial plants operating within the port

2.4.1. Cost-efficiency

Three key stakeholders prioritize economic profit: the network operator, industrial plant owners, and hydrogen producers. Their interests and concerns regarding the hydrogen network are outlined below.

Network Operator The network operator seeks to generate profit, although this may be limited by critical infrastructure regulations. The operator's expenses include network installation and maintenance, which are covered through tariff charges. Thus, setting the right tariff level is crucial for balancing costs and profitability.

Industrial Plant Owner Industrial plant owners prioritize economic gain and, therefore, advocate for lower tariffs on network usage to reduce their operational costs. Lower tariffs help enhance their profit margins by minimizing the expenses associated with accessing the hydrogen network.

Hydrogen Producers Hydrogen producers also aim for economic profit. Their profitability increases when the input costs, such as electricity or ammonia, are low, and when hydrogen prices remain relatively high in a robust market. For producers, it is particularly important that the hydrogen network connects to a broad range of potential consumers. Additionally, low tariff costs are essential to keeping the overall price of hydrogen competitive, thereby making it more appealing to potential customers.

Despite the shared goal of individual economic profit, a conflict emerges over tariff levels. The network operator benefits from higher tariffs to increase revenue, provided this does not deter customers, while hydrogen producers and consumers prefer lower tariffs to reduce costs.

From the previous analysis we observe that tariff levels are ultimately tied to the network's costs. If installation and maintenance costs are minimized, the network operator can charge lower tariffs without compromising on profit. Therefore, this thesis will concentrate on strategies to minimize hydrogen

network installation and operational costs. This approach is expected to have a positive impact on the economic outcomes for both hydrogen producers and consumers by indirectly supporting lower tariffs.

Striking a balance between cost and user facilitation is challenging, as the most cost-efficient network would involve no network at all, thereby failing to meet the needs of any users. This underscores the imperative for a decision-making method that aids in navigating the trade-off between cost and demand facilitation, while accounting for an uncertain future.

2.4.2. Facilitate users

It is crucial for the hydrogen network to effectively accommodate both existing and future users, which is a primary interest of the network operator. However, the other three stakeholders also benefit from an infrastructure that can meet current demands while being adaptable to future needs.

Future accomodation is particularly important for the port authority, as it relies on ensuring adequate infrastructure to maintain a favorable investment climate for future industrial developments (Pivetta et al., 2024). Guaranteeing sufficient capacity helps attract new industrial plants and supports the long-term growth of the hydrogen market.

2.4.3. Safety

Lastly, another requirement for a hydrogen network is its safety. Hydrogen is smaller than natural gas molecules, making it more challenging to contain in confined spaces such as storage facilities and pipelines. Additionally, hydrogen's reactivity with a wide range of materials raises further safety concerns. Hydrogen can cause embrittlement in metals, leading to the degradation of pipeline materials and increasing the risk of leaks and ruptures. This phenomenon, known as hydrogen embrittlement, occurs when hydrogen atoms diffuse into the metal, reducing its ductility and making it more susceptible to cracking under stress (Moradi & Groth, 2019). Therefore, pipelines and storage systems designed for hydrogen must be constructed from materials that resist such embrittlement to maintain their structural integrity over time.

The safety of a hydrogen network has been addressed in multiple studies (Moradi & Groth, 2019; Najjar, 2013; Weber & Papageorgiou, 2018); however, this requirement is beyond the scope of this thesis, as including it would expand the problem significantly and quickly increase computational demands.

2.5. Current state of hydrogen network planning

In practice, current energy infrastructure projects use a time horizon of 10 years to plan their hydrogen infrastructure (Cuppen et al., 2021). This planning approach minimizes installation costs of the network to facilitate the expected demand within the time horizon of 10 years. This planning approach is defined in this thesis as the immediate demand planning method; where the network is only planned for the immediate demand, namely the demand for one decade.

However, limiting the planning horizon to only 10 years may lead to repeatedly installing hydrogen pipelines on the same trajectory to facilitate the growing demand over decades. As research studies indicate that hydrogen demand will continue to grow until 2050, it can be suboptimal to only look at costs minimization and demand facilitation per decade (IEA, 2023a).

This thesis hypothesizes that the immediate demand planning method which focuses on minimizing costs within a 10-year time frame, may overlook opportunities for long-term cost reduction in the hydrogen network extending to 2055. To address this, a robust method is developed that aims to minimize costs over the entire period up to 2055 by considering the facilitation of demand up-to 2055. The robust method shall be implemented over time with an interval of 10 years; thereby considering the intervals with the end year of 2035, 2045 and 2055.

3

Theoretic Framework

The method that will be designed for this thesis will have to deal network optimization and participant and demand uncertainty. In this chapter the approach for the two aspects of the method will be discussed. It will be derived that graph theory and exploratory modeling approach will be used to develop the method. Lastly, robust and cost-efficiency shall be operationalized thereby answering sub-question 2.

3.1. Optimizing networks

To determine the most effective approach for designing a hydrogen pipeline network for an Industrial Port Cluster (IPC), a modeling technique is employed. This is necessary due to the impracticality of real-world testing. There are three primary mathematical optimization methodologies for optimizing networked systems: mixed-integer (non-)linear programming (MI(N)LP), agent-based models (ABMs), and geometric graph theory (GGT) (Heijnen et al., 2019).

Mixed-integer (non-)linear programming (MINLP) is widely used in studies of future hydrogen infrastructure (Li et al., 2019). MINLP effectively optimizes costs within energy systems, constrained by decision variables (Welder et al., 2018). Nevertheless, a notable limitation is the frequent exclusion of pipeline topology as a decision variable. Furthermore, broadening the decision space to include more flexible pipeline location options would substantially elevate the computational burden.

Agent-based models (ABMs) offer another approach. ABMs are designed for a broad range of applications and model systems from the bottom up, simulating the actions and interactions of individual agents within the system (Dam et al., 2013). Ant Colony Optimization (ACO) is a possible ABM method to optimize networks, this optimization method is inspired by ant behavior; the sources are described by ant nests and food nodes represent the sinks (Heijnen, Chappin, & Nikolic, 2014; Sivanandam & Deepa, 2008). A downside of ACO when applied to optimization networks, is that the tracks made by the ants, should be translated to a graph in order to depict it in a nice and understandable way. Furthermore, the ACO was outperformed by the GGT in computational time in a study done by Heijnen, Chappin, and Nikolic (2014).

The last approach is the Geometric Graph Theory (GGT), for this thesis GGT is chosen as approach to optimize networks. In GGT, sources and sinks are represented as nodes n (production facilities, users) and their connections (pipelines) are represented as edges e together, they form a graph $G(n, e)$ (Melese et al., 2017). Compared with MI(N)LP and ABM, graph theory is more focused on the visual presentation of networks specifically and can therefore be better used as a tool to help decision makers (Yeates et al., 2021). In addition, GGT is developed with the aim of analyzing networks and can be easily modified to meet additional system requirements, such as the prohibition of pipeline transport through specific areas, or the promotion of pipeline transport through existing pipeline corridors, something that is harder to model with MINLP (Heijnen, Ligtvoet, et al., 2014; Reuß et al., 2019).

3.1.1. Important theorems GGT

As already mentioned, GTT refers to networks as graphs $G(n, e)$ that consist of nodes n that are connected by edges e. An edge is always a connection between two nodes $e = (n_i, n_j)$. The nodes and edges can have attributes $w(e)$; for example, an edge can have the length of the edge as attribute. Graphs can have all kind of forms, for this project, we will only look at trees. A graph is a tree if there

are no cycles in the graph; meaning there is exactly one path between two arbitrary nodes. Applied to hydrogen pipeline infrastructure, a tree is the most cost-efficient way to distribute hydrogen gas 3.2. A downside, however, is that in case of a fall-out of a pipeline, some nodes will not be accessible. However, for this thesis, this will be left out of consideration.

An old graph theory problem is the minimum spanning tree (MST) problem; given a network find a tree that connects all nodes and of which the sum of the attributes of these edges is minimum (Kruskal, 1956). This problem can be solved either using Prims algorithm or Kruskals algorithm and it is known that the optimal solution is always found (Kruskal, 1956; Prim, 1957). This problem has been applied to hydrogen infrastructure before (André et al., 2013). However, as mentioned in the introduction, it is assumed that the pipelines will be placed next to the roads in the IPC. Consequently, the minimum spanning tree (MST) is not the right problem for this thesis, as the MST assumes that there are direct routes between the nodes and that they must all be connected to each other.

Therefore, in this thesis, the pipeline network is represented as a Capacitated steiner tree, which consists of two types of nodes: terminal nodes and steiner nodes. In reality, the terminal nodes represent the supply and hydrogen demand nodes and the steiner nodes represent the road or routing network. Eventually, the terminal nodes need to be connected with each other using part of the road network. The goal is to find a steiner graph that minimizes the costs while also delivering the necessary hydrogen from the supply node to the demand node. Below the minimum steiner tree and minimum capacitated steiner tree will be shortly explained.

Minimum steiner Tree

A minimum steiner tree in a graph (MStTG) is a subgraph that connects all terminal nodes, possibly incorporating some steiner nodes, while minimizing the total edge length (Winter, 1987). The task of finding a minimum steiner tree $MStT$ in a graph G is known as the minimum steiner tree in a graph problem. This problem is NP-hard, meaning there are no practical algorithm to find an exact solution in a reasonable amount of time, particularly if the size of the network increases (Žerovnik, 2015).

As a result, sub-optimal approaches, or heuristics, are used to address the minimum steiner tree problem. While these heuristics can identify a steiner tree, they do not guarantee that the solution obtained is the global minimum. Instead, they often converge on a local minimum. Numerous heuristics exist to approximate solutions to the MStTG (Winter, 1987). For instance, v. Mikulicz-Radecki et al. (2023) uses a MStTG heuristic to study the development of the hydrogen pipelines using the existing natural gas grid. They specifically applied the Minimum Spanning Tree Heuristic to identify an optimal hydrogen network integrated into the existing natural gas infrastructure (Hagberg et al., 2008).

Capacitated networks

Pipeline networks are depicted in GGT as capacitated networks, consisting of sink and source nodes. The sink represents a plant that demands hydrogen, while the source represents a supplier that provides hydrogen to the network. The capacity of a pipeline limits the volume of hydrogen that can be transported from the source to the sinks, and the network must be able to facilitate this flow. Additionally, the cost of a pipeline depends not only on its length but also on its capacity, as increased capacity requires a larger-diameter pipeline, resulting in higher material expenses. To minimize the costs of the network, it is insufficient to simply minimize the length; rather, the cost must be minimized as a function of both length and diameter. This is known as the Minimum Capacitated steiner Tree (MCStT) problem. To the authors' knowledge, there are no existing heuristics that solve the Minimum Capacitated steiner Tree problem for pipeline networks. However, there is a review of several capacitated spanning tree heuristics applied to pipeline networks by Yeates et al. (2021). Since it is possible to convert the MCStT into an MCST problem, we will first discuss the minimum-cost capacitated spanning tree (MCST) heuristics and then later adapt the chosen heuristic to a steiner tree problem (Van den Eynde et al., 2022).

Minimum-cost Capacitated Spanning Tree heuristic

Finding the MCST can be done by brute-force, where all the optimal MCST by drawing up all possible capacitated steiner tree and then selecting the one that has minimum costs. However, this brute-force approach gets extensive very fast, since a network with N nodes results in N^{N-2} possible trees (Cayley, 1857). Therefore, in practice heuristics are used to approximate the MCST.

Multiple MCST heuristics are used in the literature to plan hydrogen pipeline infrastructure. All heuristics, have the same outline; they start with an initial network (often a minimum spanning tree), capacity is then allocated and then the network is optimized with a heuristic to find a network which has less costs than the initial network (Yeates et al., 2021). In Figure 3.1, you can see an example of the different steps. Below, some network optimization heuristics will be described. This network optimization heuristic corresponds to the third arrow in Figure 3.1.

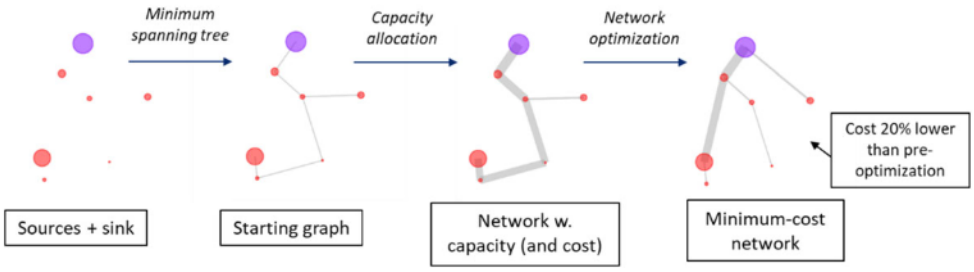


Figure 3.1: Example of the different steps of a minimum-cost capacitated spanning tree heuristic (Yeates et al., 2021)

The first heuristic is the delta change heuristic, which Rothfarb et al. (1970) introduced and André et al. (2013) implemented for hydrogen infrastructure. The essence is that cycles are created in the network with capacity by adding one edge. Then candidate networks (trees) are generated by removing one edge from the cycle, for the candidate networks, the costs are calculated. If the candidate network has lower costs than the previous network, the network is saved (Yeates et al., 2021). André et al. (2013) applied this delta change heuristic to hydrogen transmission network in France, however, to limit computational costs, only the closest node to another node was chosen to create a cycle, instead of choosing all nodes, see Figure 3.2.

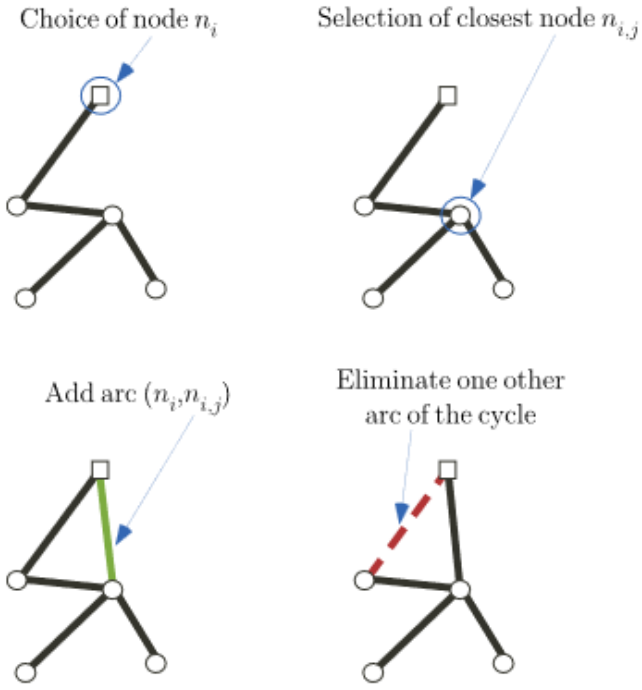


Figure 3.2: The delta change heuristic as applied by André et al. (2013)

Another heuristic is the edge turn heuristic created by Heijnen et al. (2019); which is in some sense the exact opposite of the delta change heuristic. Where the delta change heuristic creates a cycle that needs to be broken to achieve a tree. The edge turn heuristic removes one edge (n_i, n_j) of the tree thereby creating two components in the graph, see Figure 3.3. Then, candidate networks are created

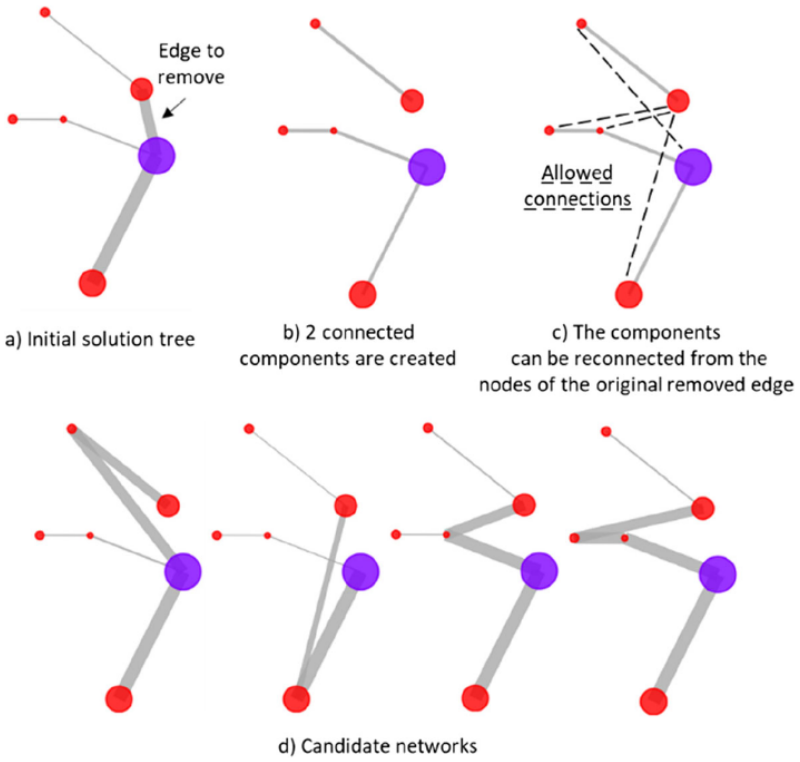


Figure 3.3: Example of implementation edge turn heuristic (Yeates et al., 2021)

by connecting the two components again with an edge that is connected with either n_i or n_j and another node of the graph n_k. These candidate networks are then optimized in such way that they can facilitate the required demand. From the candidate network and the starting network, the lowest cost network is chosen.

Yeates et al. (2021) compared different methods and found that the edge turn by Heijnen et al. (2019) had the fastest computation time compared to all heuristics, with an average computation time of 0.5 seconds for a MCST with 12 sources. The delta change heuristic, the runner up, took on average 2 seconds. A disadvantage of the edge turn heuristic is that it does not always achieve the optimal network. Nevertheless, if we aimed for significantly higher accuracy toward the optimal solution, the heuristic would get more complex, and the computation time would be almost 100 times longer (Yeates et al., 2021). These computationally complex heuristics combine a local search heuristic with a meta-heuristic. Given that we need to integrate network optimization with deep uncertainty, a fast heuristic is preferred as it will significantly decrease computational time. Therefore, the edge turn heuristic is chosen as the minimum capacitated steiner tree heuristic.

The edge turn heuristic applied to a routing network acts in essence the same. As a first step one edge is removed, creating components. Then by iterating over the nodes in the other component n_k and adding the shortest path between $n_{i/j}$ in the road network the two components are connected. Each tree is seen as a new candidate network if the shortest path between $n_{i/j}$ and n_k is not partly located in one of the components, see Figure 3.4. Then the candidate network is selected with the lowest costs.

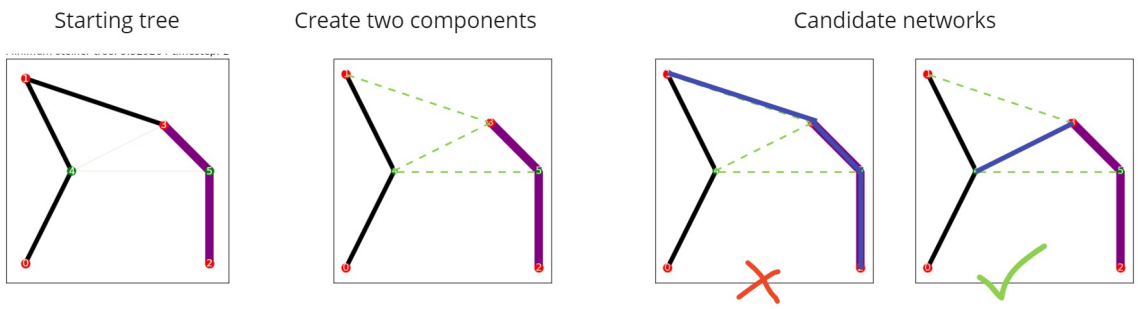


Figure 3.4: Example of the edge turn heuristic applied to a routing network

3.2. Decision Making under Deep Uncertainty

The network optimization method that will be developed using GGT will be used as a decision making support tool. However, making decision about the hydrogen infrastructure of an Industrial Port Cluster is very difficult as the hydrogen demand is deeply uncertain due the amount of external factors that influence this hydrogen demand. Deep uncertainty means that because of the complexity of the system, there are a large number of plausible futures that cannot be ranked based on the likelihood of their occurrence (Maier et al., 2016).

To model plausible futures for a deep uncertain future scenarios should be generated (Herman et al., 2015). There are three kind of scenarios; predictive scenarios, exploratory scenarios and normative scenarios. First kind of scenario is a predictive scenario. A predictive scenario can also be called a what-if or a best-guess scenario (Maier et al., 2016). Predictive scenarios are used if the system is relatively well-known and well-defined (van Vuuren et al., 2012). Another kind of scenario modeling is with exploratory modeling; where the multiple future space is explored to answer the question *what could happen* (Maier et al., 2016). The last kind of scenarios that can be used to model future scenarios is the normative scenario; with normative scenarios generation the model is more solution focused and tries to answer a question like how can a specific future be realized?

For this thesis, exploratory scenario generation will be employed because the hydrogen energy system is still in its early stages and not yet well-defined, making predictive scenarios unsuitable (Paredes-Vergara et al., 2024). Normative scenarios are also inappropriate, as there is no single, shared future that the group of stakeholders aims to achieve independently of external factors. Therefore, this thesis will utilize exploratory modeling to examine how external factors beyond the influence of the stakeholders affect the future of the network (Börjeson et al., 2006). The Exploratory Modeling Analysis python package developed by J. H. Kwakkel and Pruyt (2013) is used in this thesis to generate the exploratory scenarios.

This thesis uses Many-Objective Robust Optimization (MORO) as robust optimization approach as it uses exploratory scenarios to optimize the robust solution, effectively tackling uncertainty from the beginning (Bartholomew & Kwakkel, 2020). In contrast to other robust decision making methods, which only use several or one reference case for robust optimization. MORO measures the robustness for each outcome by sampling from the plausible future space, guaranteeing that solutions function satisfactorily across all plausible futures.

Eventually, the method will support decision makers in developing a robust design for the hydrogen pipeline network. The robustness definition as practiced by this thesis is that robustness is a measure of the *adequate performance* of the design under a range of plausible scenarios (Maier et al., 2016). However, the precise definition of adequate performance is not globally determined, on the contrary the robustness measure should be defined by the interested parties. An appropriate robustness metric is determined by the combination of the likely impact of system failure and the degree of risk aversion of the decision-maker (McPhail et al., 2018).

3.3. Synthesis Theoretic Framework and Problem definition

In this section, the problem definition and the theoretic framework shall be synthesized and summarized, thereby defining cost-efficiency and robustness.

Since the hydrogen pipeline network method should be a decision making support tool, a geometric graph theory approach is used to optimize networks as this is a intuitive and visual approach to optimize networks and it makes it easy to add additional constraints or nodes. As the industrial port cluster is densely built, it is assumed that the pipelines will follow the road network. Thus, the optimization pipeline network problem is a capacitated minimum steiner tree problem. Using the actor analysis, the objective of the network is to minimize costs and to facilitate demand over time. To address the over time aspect the method shall optimize the hydrogen network every ten years (2035, 2045, and 2055), it thereby represents the standard monetary planning horizon of infrastructure development. Furthermore, it will also be able to depict the path-dependency that is attributed to pipeline network infrastructure by using the developed network of the previous years as input for the next ten years.

Due to the uncertainty in industrial hydrogen demand towards 2050, it is uncertain how much demand should be facilitated over-time. To take demand and participant uncertainty into account exploratory scenario generation will be used for multi-objective robust optimization, by optimizing networks for a lot of plausible future demands. The goal of generating all these networks is to create a robust network from all these scenarios, that optimizes a robustness measure related to the facilitation of demand. A robustness measure represents the risk-adverseness and values of stakeholders and should therefore be defined by the decision makers.

In this thesis, it is assumed that the robustness measure should portray a high risk aversion for missed facilitated demand. This assumption is based on the study of Cuppen et al. (2021), where a collaborative approach was adopted to investigate possible interventions in the energy infrastructure with important stakeholders. Stakeholders that were actively involved in that study were the network operators, the industrial plant owners, and the port authority. The collaborative robustness metric defined in that study was a negative definition of success: the least number of missed transition events from the pathway of events. This corresponds with a high risk aversion for missed facilitation of demand; this stance of decision makers is what we will implement for this research as well. The robustness measure evaluates the performance value of facilitated demand, which we aim to maximize.

Since we want to optimize the facilitated demand, a regret-based metric is preferable, especially when combined with cost minimization (McPhail et al., 2018). High risk aversion metrics that account for regret include minimax regret and the 90th percentile (P_{90}) minimax regret. From these two, the minimax regret exhibits the highest degree of risk aversion in comparison to the 90th percentile minimax regret. For this thesis, the 90th percentile minimax is selected, as it displays reduced sensitivity to outliers.

4

Methodology

The current hydrogen infrastructure planning method is the Immediate Demand Planning Method (IDPM), as described in Section 2.5. It uses a 10 year time horizon to plan cost-efficient infrastructure, which could lead to increased costs over time. This chapter describes a new method for developing a robust and cost-efficient network within an IPC, which aims to minimize costs over 30 years. This method is called the Robust Backtracking Planning Method (RBPM).

This chapter is organized into several sections, with a visual summary provided in Figure 4.1. First, Section 4.1 introduces the IPC model, which is based on the observed IPC characteristics from Chapter 2. This model will later be used to test the different planning methods. Next, Section 4.2 explains the process of generating demand scenarios, using the IPC model as input. In Section 4.3, it is described how the networks are optimized towards 2055 from the current infrastructure planning method point of view; per decade and per demand scenario. In Section 4.4, it is reported how these optimized networks can be visually analyzed to better understand the future solution space of the IDPM. Following this, Section 4.5 presents two methods for generating robust networks. Finally, Section 4.6 describes the implementation of these robust methods over time.

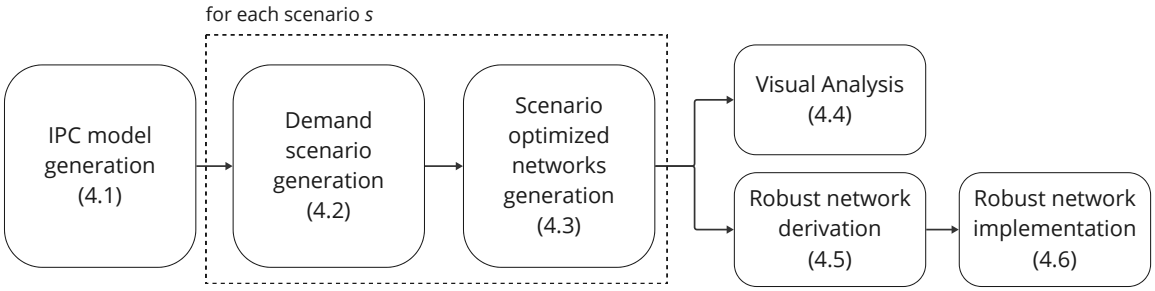


Figure 4.1: Visual overview of the methodology, each blocks represents a flow block in the method and the corresponding section is indicated

4.1. IPC Model

This thesis is not using a case study, like an existing industrial port cluster. Instead, for this study small industrial port clusters are generated. The purpose of generation the IPC model is first of all to illustrate the method (Edmonds, 2017). Secondly, the models goal is a theoretical exposition where the model is used to test different hypotheses of the robust method regarding different clusters.

In this IPC model the main characteristic derived in chapter 2 will be incorporated. The main characteristics of an IPC related to hydrogen networks are

- High presence of water ways and docks that have a tree like structure
- A road network that is heavily influences by the water ways, that does not often cross the water, and its road network is structured around the water
- High concentration of firms
- Large share of industrial plants from the same industrial subsector
- Hydrogen network follows the road network due to presence of water and high concentration of firms.

For modeling purposes two concepts are distinguished; the IPC layout and the IPC configuration. An IPC layout consists of waterways, road network, and plant locations. An IPC configuration is a combination of the IPC layout and the assignment of plant specifics to the plant locations. Below, first the port layout generation step is discussed, and then the IPC configuration step is presented.

4.1.1. IPC layout generation

The output of this step is a port layout with industrial nodes, a road network, and the wet infrastructure. The port layout is generated using a method developed by Petra Heijnen to generate case studies for project groups in the course *Prestatie Analayse in Energie en Industrie* for the Technology, Policy and Management bachelor program. An overview of the port layout generation model can be seen in Figure 4.2.

In Step 1, waterways are created by generating a minimum spanning tree (MST) from randomly generated locations within 80% of the total dimensions of the port. These randomly generated locations serve as the waterway points. This MST is subsequently connected to a rectangle that represents the water at the bottom of the graph. From the MST, a polygon is formed by adding width to the edges and smoothing the lines by removing polygon points that are less than ϵ distance apart. This process yields a set of randomly located waterways. The author conducts a visual inspection of these waterways to ensure they resemble the wet infrastructure described in Section 2.1.2. The primary visual assessment criterion for the waterways is that they should effectively represent a dock structure.

In Step 2, a road grid is generated by overlaying a grid of squares across the entire harbor layout, with each grid line corresponding to the specified resolution. Grid points located within or near a 20-meter radius of the waterways are subsequently removed. Additionally, grid corners are randomly eliminated with a 10% probability to create a more varied composition of the road grid in areas unaffected by water. The coordinates of the road grid are further adjusted randomly within a distance of 5 meters in both the x and y directions. Given that a strictly grid-based infrastructure would normally favor the shortest path routes, this small change in location is required to avoid the formation of uniform paths.

In Step 3, the locations of industrial plants are assigned to nodes within the road grid. Each industrial node is preferably situated at the end of a road, which is defined as a node with only one neighboring connection. In cases where there are insufficient endpoints available, the remaining industrial locations are placed arbitrarily at other points within the road grid. Throughout the placement process, the author visually assessed the distribution of industrial nodes, ensuring that at least one industrial node is assigned to each quarter of the IPC area to achieve a balanced representation.

In Step 4, a subset of the road grid is selected to form the road network. In reality, the road network is primarily developed to facilitate the efficient transport of produce to other locations, as discussed in section 2.1.2. Consequently, the road network is derived from the road grid by identifying the roads that lie along the shortest paths between two arbitrary industrial nodes or those that are part of the minimum steiner tree that connects all industrial nodes. Furthermore, nodes that Lastly, to simplify the

road network, nodes located on straight roads that have a degree of 2 are eliminated, as they do not contribute additional value to the road network.

The final road network $H(N, E)$ is a steiner graph with the industrial nodes as terminal nodes $N_{\text{terminal}} \in N$ and the grid corners are represented by the steiner nodes $N_{\text{steiner}} \in N$.

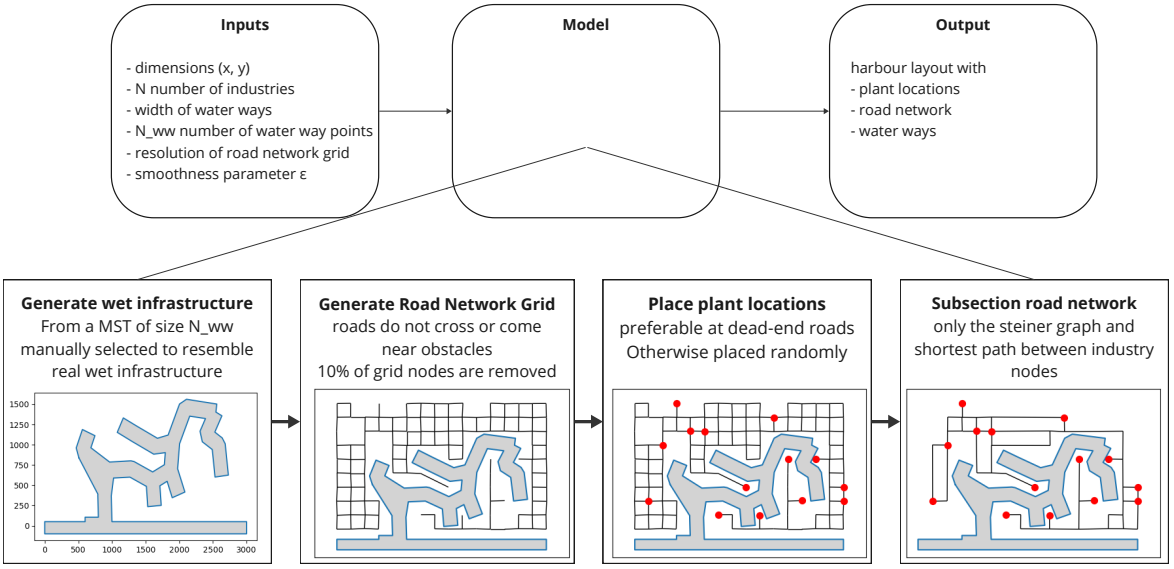


Figure 4.2: Schematic overview of the harbour layout generation model

It is imperative to acknowledge that the algorithm for generating waterways is somewhat random and may not accurately reflect real-life conditions. For example, the real life wet infrastructure of a port does not conform to a minimum spanning tree. Unfortunately, due to time constraints, it was not feasible to develop a more realistic algorithm for generating the wet infrastructure of the port layout.

4.1.2. IPC configurations

For one IPC layout multiple configurations can be generated. A configuration is generated from the IPC layout by assigning the industrial nodes as either a supply node, or a demand node.

Supply node

In each configuration, a single supply node represents the hydrogen production and storage facilities in the IPC. Initial tests revealed that the location of this supply node significantly impacts network outcomes. Additional research revealed that facilities for producing hydrogen are intended to be situated in close proximity to one another (CES Rotterdam-Moerdijk, 2021). As a result, the decision was made to display the supply facilities as a single network node.

Additionally, the supply node is restricted to the lower quadrant of the IPC, near major water bodies, to align with regulations that position flammable materials away from populated areas (Koningsveld et al., 2023). Since this supply node represents both hydrogen production and storage, it is assumed to meet all demand requirements. Without this assumption, supply variability would need to be accounted for, potentially fluctuating throughout the day (Caglayan et al., 2021). This variability in supply, however, falls outside the scope of this thesis, which assumes a continuous and sufficient hydrogen supply from the designated supply node.

Demand node

Demand nodes are conceptual representations of existing industrial plants associated with a specific process within a subsector. For simplicity, each demand node is modeled to feature only one process that can transition between different states. This simplification is justified by the relatively small scale of the IPC; if a company has multiple potential hydrogen processes, these are represented in the IPC as separate nodes.

Table 4.1 outlines the properties assigned to each demand node and describes their interrelationships. In addition to the subsector, process, and decarbonization technology, each demand node is assigned a production capacity, which indicates the amount of product produced annually.

Furthermore, as this thesis posits that the likelihood of transitioning to hydrogen processes is influenced by the specifics of the processes and technologies involved. Consequently, each demand node is categorized by its process temperature and assigned a Technology Readiness Level (TRL). Finally, to project future hydrogen demand, each demand node is also associated with a Specific Energy Coefficient (SEC), which quantifies the energy consumption of the decarbonization technology per ton of output product.

Some demand nodes are labeled as *extra*. They represent the possibility of another industrial plant relocating to this particular IPC or an existing plant expanding its operations. In future demand scenarios, the likelihood of an extra node appearing will reflect the business climate in the IPC, which is influenced by a variety of external factors.

Table 4.1: Table to list the properties given to each industrial node in a configuration

Property	Assignment criteria	Unit
Subsector	Sub-sector picked from the subsector list identified as likely candidates to transition to hydrogen as decarbonization strategy	-
Process	Given the sub-sector, a corresponding process is randomly picked	-
H_2 decarbonization technology	Given the process, a corresponding hydrogen technology is picked as decarbonization technology	-
Production Capacity	Given the process, the production capacity is picked from a normal distribution	ton/year
Process Temperature Category	Attribute to the Process, given the process, the associated temperature category based on Table 2.1	
Specific Energy Coefficient	Attribute to he H_2 decarbonization technology	
Technology Readiness Level	Attribute to the H_2 decarbonization technology	

4.1.3. Cluster generation

As previously defined in Chapter 2, IPCs consist of geographical close companies operating within the same specialization field, such as a specific subsector. The IPC configurations are conceptualized to represent a cluster in which at least c_c percent of the industrial nodes belong to a single subsector. For instance, in the context of this thesis, we refer to an Iron and Steel cluster if at least c_c % of the plants within that IPC are classified as Iron and Steel industrial plants.

4.2. Scenario generation

Given an IPC configuration many plausible future scenarios will be generated in the demand scenario generation phase using EMA work package (J. H. Kwakkel & Pruyt, 2013), a schematic overview of this step can be seen in Figure 4.3. The inputs of this step are the plant specifics from the IPC configuration, the uncertainty parameters and the technology specific energy data.

Each demand scenario specifies hydrogen demand for every industrial plant in the years 2035, 2045, and 2055. These scenarios are generated based on the transition probability $p_{n,t}$ for each plant node n at timestep t and the estimated hydrogen demand $D_{\text{est},n}$ if the plant transitions to hydrogen. At each timestep, the plant transitions to hydrogen with probability $p_{n,t}$ (see Equation 4.1). Notably, if the technology readiness level $TRL_n \leq 6$, the probability of transition in 2035 is set to zero. The flow diagram illustrating the generation of a single demand scenario is provided in Appendix Figure D.1.

$$D_{n,t} = \begin{cases} D_{\text{est},n} & \text{if } p < p_{n,t} \text{ or } D_{n,t-1} = D_{\text{est},n}, \\ 0 & \text{otherwise,} \end{cases} \quad \forall n \in N_{\text{demand}},\ p_{n,t} \in [0,1],\ t \in \{2035, 2045, 2055\} \tag{4.1}$$

A simple example of a demand scenario can be seen in Figure 4.4. As one can notice, the hydrogen demand per node is either on or off there is no gradual aspect on this function. This is done since one node resembles one process, and we will assume that this full process will transition to hydrogen. However, in reality it is can be the case that only a small portion is first transitioned as pilot project or that the plant will expand. This subtlety is not modeled in these scenarios.

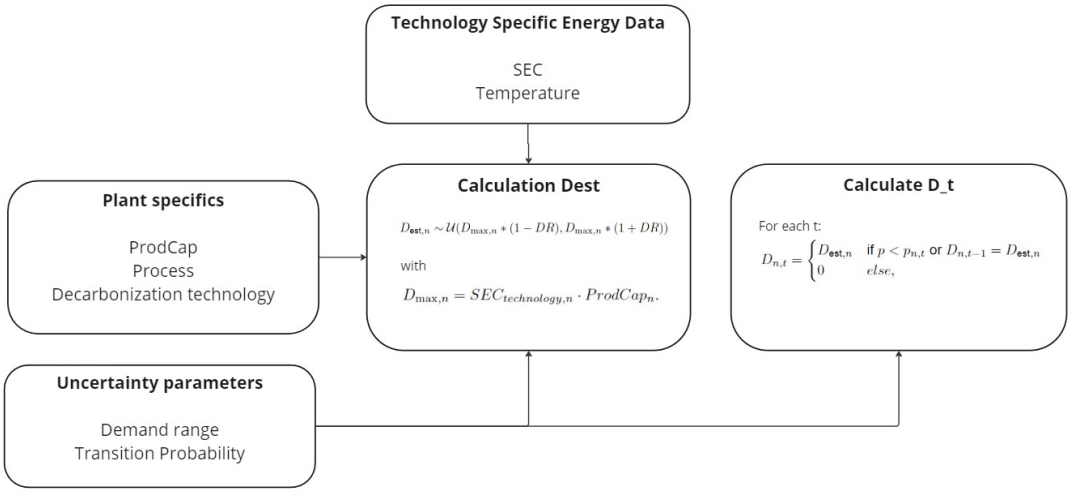


Figure 4.3: Overview of the demand generation of one node for one scenario

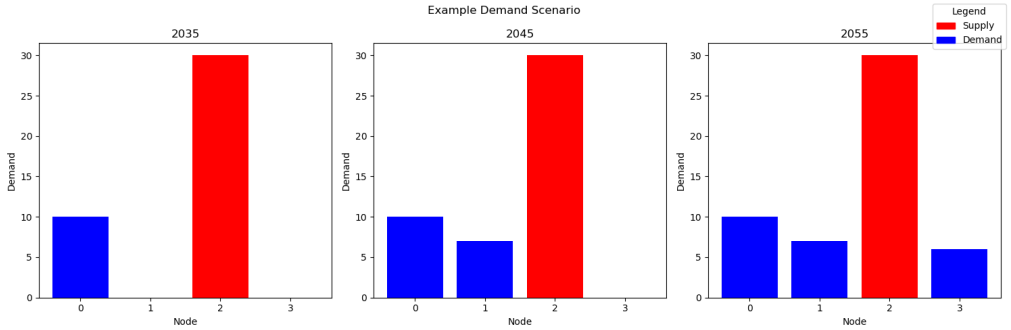


Figure 4.4: An example of one demand scenario. One demand scenario contains the future demand and supply for the demand nodes in the years 2035, 2045 and 2055.

4.2.1. Hydrogen demand per process

The estimated maximum hydrogen demand $D_{\text{est},n}$ is based on Neuwirth et al. (2022a). In their paper, the maximum hydrogen potential per industrial plant $D_{\max,n}$ is determined using the specific energy consumption (SEC) and the plant specifics. SEC represents the energy use of the technology per ton of output product. Thereby, if the process would transition from fossil fuel use to hydrogen use, the maximum hydrogen demand $D_{\max,n}$ is a function of the production capacity $ProdCap_n$ and the Specific Energy Consumption of the technology $SEC_{technology,n}$:

$$D_{\max,n} = SEC_{technology,n} \cdot ProdCap_n. \tag{4.2}$$

with maximum hydrogen demand $D_{\max,n}$ in TWh/year, specific energy consumption $SEC_{process,n}$ in MWh/ton, and production capacity $ProdCap_n$ in Mton/year.

The approach of estimating hydrogen consumption based on current industrial process characteristics is also adopted by Namazifard et al. (2024), who explored several potential hydrogen supply scenarios for Belgium in 2050. However, while their focus is on supply, this thesis emphasizes the hydrogen network as a function of demand. In different future scenarios, the hydrogen required for transitioning industrial processes may vary due to technological advancements, as well as slight expansions or reductions in plant capacity during the transition. To account for this uncertainty in the future hydrogen demand, we model the estimated demand at each node, $D_{\text{est},n}$, as a uniform distribution. This distribution accounts for possible deviations from the maximum demand $D_{\max,n}$ per scenario. Specifically, the estimated demand is assumed to lie within a range defined by the demand range Δ_D. The uniform

distribution is given by:

$$D_{\text{est},n} \sim \mathcal{U}(D_{\max,n} \cdot (1 - \Delta_D), D_{\max,n} \cdot (1 + \Delta_D)) \tag{4.3}$$

where $D_{est,n}$ is uniformly distributed between $D_{\max,n} * (1 - \Delta_D)$ and $D_{\max,n} * (1 + \Delta_D)$. The demand range Δ_D is a number between 0 and 1, and $D_{\max,n}$ is the calculated maximum potential demand given in equation 4.2.

This approach reflects the assumption that future demand could vary symmetrically within this interval, accounting for both potential increases and decreases in demand due to factors such as technological improvements or capacity changes. It is important to note that once a node undergoes the transition, its hydrogen demand is assumed to stabilize and remain constant.

4.2.2. Uncertainty Parameters

Multiple scenarios are generated using the Exploratory Modeling Analysis (EMA) work package (J. H. Kwakkel & Pruyt, 2013). As explained in section 3.2, with EMA it is possible to generate exploratory scenarios by formulating uncertainty parameters and the uncertainty parameter space. The EMA work package then, sample from this uncertainty parameter space using Latin Hypercube Sampling (LHS). LHS splits the uncertainty space by the number of scenarios and randomly combines these values , this way a unique and uniform distribution of values across scenarios is ensures (Huntington & Lyrintzis, 1998).

The uncertainty parameters used in this thesis can be seen with their definition in Table 4.2. We distinguish three types of uncertainty; the transition uncertainty p_T, the settle uncertainty p_{extra} and the demand uncertainty Δ_D.

Firstly, the probability of an industrial plant transitioning to hydrogen decarbonization technology is characterized as an uncertain parameter, as it relies on numerous external factors. These external factors are encapsulated in the transition uncertainty p_T, as referenced in equation 4.1. In this thesis, three transition probabilities p_T are distinguished: low, medium, and high categories. These categories correspond to the likelihood of a plant successfully transitioning. As discussed in Chapter 2, processes that require high-temperature heat or serve as feedstock have a greater likelihood of transitioning compared to those that operate with low-temperature heat. The assignment of the transition probability category is based on the temperature category of each process.

Secondly, another uncertainty parameter is chosen; p_{extra} the probability of an extra node occurring in the IPC. When this node appears in the IPC, then it will be treated just as a normal demand node with each decade a chance of p_T to transition. Thirdly, the last uncertainty parameter is the demand uncertainty, which is represented as the demand range Δ_D, see equation 4.3.

Table 4.2: Uncertainty parameters for EMA used to generate demand scenarios

Uncertainty Parameter	Definition
p_{lowT}	The chance that a plant transition with low temperature
p_{midT}	The chance that a plant transition with medium temperature
p_{highT}	The chance that a plant transition with high temperature or with a feedstock process
p_{extra}	The chance that an extra industrial plant appears
Δ_D	The range of which the estimated demand $D_{\text{est},n}$ deviates from $D_{\max,n}$

4.3. Network optimization

This network optimization step should resembles the immediate demand planning method (IDPM), where a time horizon is executed of one decade. For each decade, a network is identified that meets the demand for that specific decade while minimizing installation costs.

4.3.1. Optimal Network Layout Tool

The network optimization for one decade is based on the Optimal Network Layout Tool (ONLT) developed by Heijnen (2024). The ONLT is a graph-theory-based tool that employs the edge turn heuristic to identify a minimum capacitated steiner tree. This tool is highly flexible regarding its inputs, as it is

designed for easy change in inputs such as demand, nodes, road networks, and existing infrastructure, thereby assisting in multi-actor decision-making processes. This input flexibility makes the ONLT particularly suitable for this thesis, which aims to analyze various IPC layouts under different demands.

The ONLT is designed to find one cost-efficient network based on multiple demand scenarios. In the original ONLT framework, a single node can function as either a demand or supply node at different time steps and the demand can increase and decrease. However, in this thesis, nodes are distinctly classified as either demand or supply nodes. This distinction is due to the small scale of the IPC, where it is assumed that industrial plants do not return hydrogen to the system. Additionally, if the supply amount and location is constant the ONLT will always solve its network for the scenario with the highest demand. Since the demand in this study is only anticipated to increase over time, the ONLT will only be given one demand scenario, which is the scenario at the end of the decade, which naturally has the highest demand of the entire decade.

Additionally, for this thesis, the ONLT is adapted for use across three decades 2025-2035, 2035-2045, 2045-2055 indicated by the end year: 2035, 2045, and 2055. The ONLT finds the minimum capacitated steiner tree for one timestep so for each scenario the ONLT is ran three times resulting in three capacitated graphs $G_{2035}, G_{2045}, G_{2055}$ where G_{t-10} is an input for the calculation of G_t. This way, the path dependency of installing pipelines using the IDPM and the influence this has on the future network is modeled.

This adaptation allows the ONLT to represent the Immediate Demand Planning Method (IDPM) with three times a time horizon of one decade. At each time step, the ONLT minimizes the costs for installing new pipelines that are needed to facilitate demand.

4.3.2. Cost function

The cost function is used to optimize each scenario network, therefore the cost function is an important part of the network optimization. The cost function of the ONLT is a unitless cost function as the ONLT is not specifically tailored to optimize hydrogen networks, but rather any commodity in pipelines, such as natural gas or biogas (Heijnen, Ligtvoet, et al., 2014).

For this thesis, the ONLT cost function is substituted with a cost function that is already applied to a local hydrogen infrastructure network by Hammond et al. (2024). This function takes into account the labor, material and other various costs needed to install the pipelines. Originally, Hammond et al. (2024) made a cost function in pounds, therefore, for this thesis the cost function is conversed to million euros using the average conversion rate of 2024 (X-Rates, 2024).

The cost function $C(t)$ consists of capital costs $C^C(t)$ and operating costs $C^O(t)$. The operational costs are taken as 4% of the total capital costs per year (Hammond et al., 2024). For the edges (i, j) that are newly installed in decade t, the operating costs are assumed to only be applied for 5 years. This assumption is made as some edges will be installed early in the decade, and other later, so on average it is set that the operational costs of the installed edges at timestep t are halved.

$$C(t) = C^C(t) + C^O(t) \tag{4.4}$$

$$C^O(t) = (0.04 \cdot 10) \cdot \sum_{i=2035}^{t-10} C^C(i) + (0.04 \cdot 5) \cdot C^C(t) \tag{4.5}$$

$$\tag{4.6}$$

where $C(t)$ is the total costs in million euros for one timestep t, $C^C(t)$ the capital costs in million euros for one timestep, and $C^O(t)$ the operational costs in million euros for one timestep.

The capital costs $C^C(t)$ at timestep t is the costs for the installation of newly installed edges. Newly installed edges are portrayed as the difference between the network at the current timestep G_t and the already installed network G_{t-10} if it exist. This includes edges that are not placed yet or edges that require a higher capacity.

$$C^C(t) = \sum_{(i,j)\in G_t \setminus G_{t-10}} c_{ij}^C \tag{4.7}$$

$$c_{ij}^C(l_{ij}, d_{ij}) = 50833 l_{ij} e^{0.0697 d_{ij}} + 297 (d_{ij})^2 + 71800 d_{ij} + 546582 \tag{4.8}$$

where c_{ij}^C is the capital costs of one edge between nodes i, j with length l_{ij} in meters and outside diameter d_{ij} in meters.

The ONLT can only influence the edges (i, j) installed at timestep t, as the edges installed in previous decades are already installed. Therefore, the ONLT aims to minimize $C^C(t)$ in equation 4.7. The ONLT is designed to alter both the length l_{ij} and the capacity q_{ij} of an edge to look for cost minimization. However, the selected cost function is based on the diameter d_{ij} rather than the capacity, which is the primary focus of the ONLT method. To determine the relationship between diameter d_{ij} and capacity q_{ij}., a simplification of the fluid flow equation is conducted. For a detailed derivation, please see appendix C.1. The simplified formula for the diameter d_{ij} is as follows:

$$d_{ij}(l_{ij}, q_{ij}) = \left[0.27\ q_{ij}\ l_{ij}^{0.5}\ \right]^{-2.5} \tag{4.9}$$

where l_{ij} is the length of the pipeline in meters, and q_{ij} the capacity of the pipeline in ton H_2/day.

In Figure 4.5 the behavior of the capital costs C^C as function of the length and capacity can be seen. As one can see the costs scale almost linearly with l_{ij}, in the pipeline length range of an IPC $[0, 1$ km$]$. The behaviour of C_{ij} as a function of l_{ij} corresponds with the pipeline cost function defined in the ONLT where $c_{ij} = l_{ij} * q_{ij}^{\beta}$ (Heijnen, Ligtvoet, et al., 2014). Looking at the behaviour of c_{ij}^C as a function of the capacity, one can see that when q_{ij} gets large, the dependency with the costs assumes a linear form. Depending on the length of the pipeline, the capacity has more influence on the rising costs. This is because when l_{ij} is larger, the term amplifying $e^{0.0697 d_{ij}}$ also grows thereby amplifying the effect of the capacity on the costs.

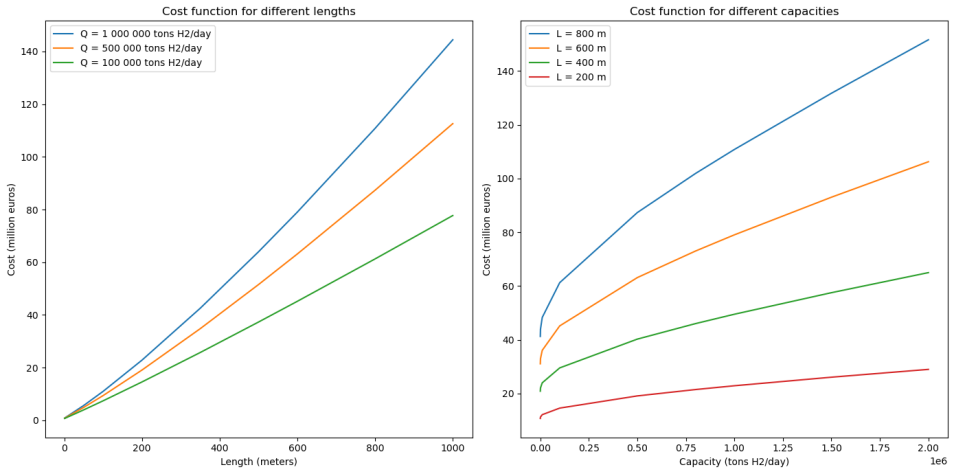


Figure 4.5: Plot to visualize behavior of the cost of one edge as a function of the edges length and capacity.

4.3.3. Network optimization algorithm

In this subsection the algorithm is described that is used to minimize the capital costs (Eq. 4.7) every decade. The inputs of this network optimization algorithm can be seen as an example in Figure 4.6 and are listed below.

- Road network $H(N, E)$ with the industrial nodes as terminal nodes $N_{\text{terminal}} \subset N$
- Demand $D_{s,t,n}$ for each scenario s for each timestep t and terminal node $n \in N_{\text{terminal}}$
- Network of already placed pipelines in previous timestep G_{t-10}. This network is empty for G_{2025}

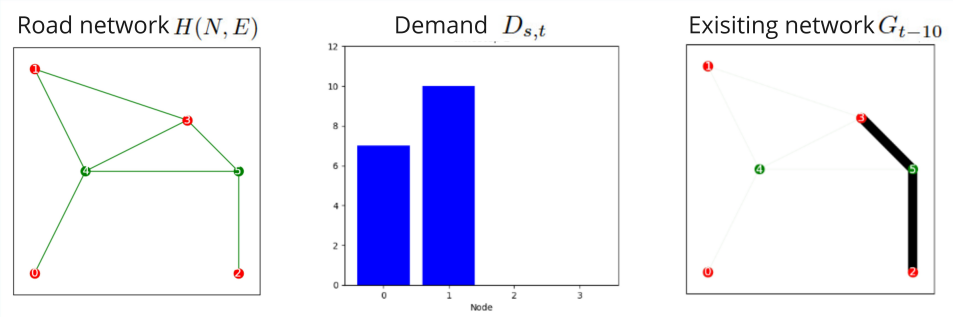


Figure 4.6: Example of inputs needed for the optimization network algorithm for one timestep

Using these inputs a capacitated network is generated that minimizes costs while facilitating the demand of $D_{s,t}$. This happens in two steps, these steps are depicted in Figure 4.7, and are further described below.

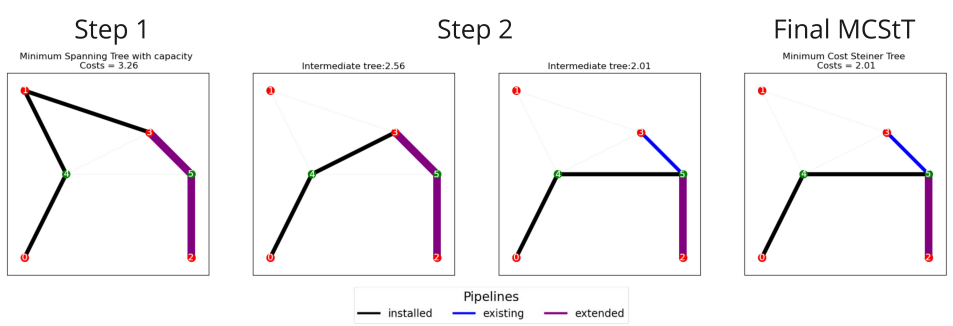


Figure 4.7: Flow of the ONLT for a steiner Tree in a Graph problem for one timestep

Step 1 **Generate Minimum Steiner Tree:** The minimum steiner tree (MStT) is determined using the distance network heuristic developed by Kou et al. (1981). Using this heuristic, first a complete graph between all the demanding terminals, where the edges resemble the shortest path in the steiner graph is generated where the existing edges are used as existing connections. Then from this complete graph the MST is calculated that spans the active terminal nodes. Lastly, the complete graph MST is translated back to the road network by replacing the edges by the shortest paths in the steiner graph. The capacities are assigned to the edges of the MStT is such way that the demand of the nodes can be delivered from the supply node to all the demand nodes. Then the network costs as a function of the edges capacities q_e and the edges length l_e is calculated. The result is an approximated minimum length steiner tree with capacity and a certain costs

Step 2 **Search for (near) Minimal Capacitated Steiner Tree:** A network with lower cost is search for by removing each new pipeline, starting with the longest, and with the edge turn heuristic look for a new one (recall Section 3.1). The objective here is the least cost, and the constraint is that all demand should be facilitated for this timestep. The already existing pipelines are not removed. This heuristic is iterated until all edges are visited.

4.3.4. Example of outcome

The final output of the network optimization algorithm are three networks G_{2035}, G_{2045}, and G_{2055} which belong to the same demand scenario D_s. An example of an outcome is depicted in Figure 4.8. Where the green edges indicate the road network, the black edges the newly installed pipelines, the blue edges the previously installed pipelines, and purple edges indicates an extension of previously installed pipelines. One can see that the pipelines between the supply node and node 5 is extended every decade. Since every decade the hydrogen demand increases and as a consequence that pipeline needs to be extended.

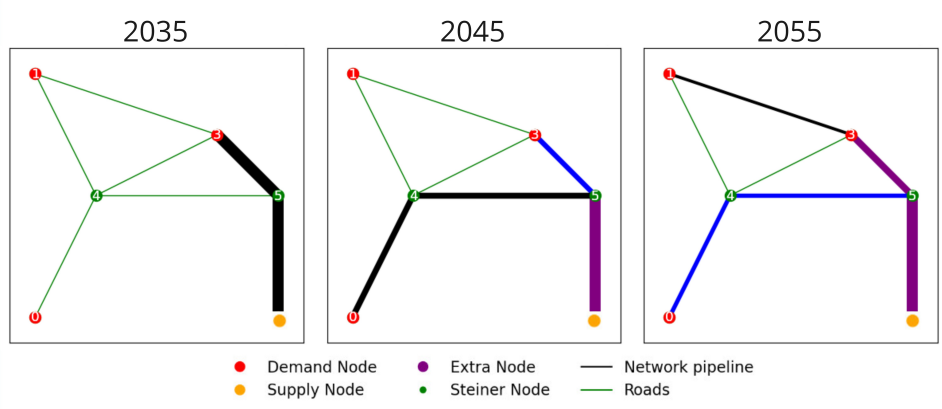


Figure 4.8: Example of the outcome of the network optimization; three networks for 2035, 2045, and 2055.

4.4. Analysis scenario-optimized networks

There are multiple scenario-optimized networks, each representing a potential future reality; however, the likelihood of any specific scenario happening is almost null. Consequently, during the analysis phase, the scenario-optimized network structures are visually examined to identify emerging networks and patterns.

Two visualization tools are utilized in this methodology: the overlapping capacity graph and the occurrence graph. These visualizations are adapted from the plotting functions of the ONLT, which already includes several features for effectively plotting network data. The plotting function of the ONLT can plot one network in the IPC area, indicating the terminal nodes and the steiner nodes, and the capacity of each edge in that network is indicated by the thickness of the edge.

The overlapping capacity graph shows the capacity of an edge over all scenarios. The thickness of an edge corresponds to its capacity, while the darkness of the edge represents the frequency with which that capacity is observed across all scenarios.

The occurrence graph visualizes the frequency with which an edge appears across the demand scenarios, specifically showing how often edge $(i, j) \in N_t$ occurs. This is represented in an IPC network plot, where the thickness of an edge reflects the number of times that edge is counted in all the demand scenarios.

This visual analysis is particularly useful during the exploratory phase of a multi-actor decision-making process, as it can illuminate the possible outcomes of the network (Heijnen, Ligtvoet, et al., 2014). Furthermore, the occurrence and overlapping capacity graphs contribute to the interpretation of the robust network, the development of which will be discussed in the next section.

4.4.1. Facilitated Demand

Furthermore, these scenario-optimized networks are also analyzed by looking at their performance in the whole future demand scenario space. The cost function is already discussed in the previous section, but here the facilitated demand objective will be described and its associated robustness measure, after which the multi-objective optimization problem at hand will be discussed.

The facilitated demand is an important objective for the stakeholders involved with planning a hydrogen network. The facilitated demand of a network in one scenario $FD_{G_{i,t},s,t}$ is defined as the maximum flow Φ the network can facilitate for that demand scenario;

$$FD_{G_{i,t},s,t} = \Phi_{G_{i,t},D_{s,t}}. \tag{4.10}$$

Where $FD_{G_{i,t},s,t}$ is the facilitated demand in ton H_2/day for network i at timestep t, and $\Phi_{G_{i,t},D_{s,t}}$ the maximum flow the network can facilitate for that specific demand scenario $D_{s,t}$ in ton H_2/kg.

The range of $FD_{G_{i,t},s,t}$ extends from 0 to $D_{s,t}$, with $D_{s,t}$ representing the optimal value when all demand is satisfied, and 0 indicating the scenario where no demand is met. The decision to express the facilitated demand in tons H_2/day rather than as a unitless fraction based on the percentage of the maximum facilitated demand is due to the facilitated demands' correlation with potential profit. From a profitability standpoint, it is significant whether 2% of 1000 or 2% of 20 is not facilitated. Moreover, from a decarbonization perspective, the total quantity is also more significant than the proportion; an increased volume of H_2 per day correlates with a greater production of decarbonized products.

4.4.2. Robustness Measure

In order to compare the networks over the different scenarios, we are specifically interested in the facilitated demand over the set of plausible futures. As discussed in section 3.2, in principle the robustness measure should be chosen in consultation with the decision makers. However, as this thesis develops a generic method, thus, an example robustness measure is determined. During application of this method to a real case example, the robustness metric should be determined in cooperation with actors. In this thesis, it is determined that the actors value high risk aversion for missed facilitated demand, as already mentioned in Section 3.3.

High risk aversion metrics that account for regret include (100th percentile) minimax regret and the 90th percentile (P_{90}) minimax regret. Where the minimax regret exhibits the highest degree of risk aversion in comparison to the 90th percentile minimax regret. For this thesis, the 90th percentile minimax is selected, as it displays reduced sensitivity to outliers. Below first the calculation of the minimax regret shall be described, and then the 90th percentile.

The minimax regret $MR_{i,s,t}$ is a relative performance value as it represents the regret one has in a certain scenario compared to the calculated optimal outcome of that scenario (Savage, 1951). Applied to our case, the maximum regret is the facilitated demand a network $G_{i,t}$ has for one scenario s, compared to the maximum facilitated demand all networks $G_{j,t}$ can achieve for the same scenario;

$$MR_{G_{i,t}}(s) = \max_{j}(FD_{G_{j,t},s,t}) - FD_{G_{i,t},s,t} \tag{4.11}$$

Where $MR_{G_{i,}}(s)$ is the maximum regret in ton H_2 /day for one scenario at one timestep. The optimal value of $MR_{G_{i,t}}(s)$ is 0. This equation can be simplified if each network $G_{i=s}$ is optimized for scenario s as is done with the ONLT tool. If so, the term $\max_j(FD_{G_{j,t},s,t})$ is always equal to the $FD_{G_{s,t},s,t}$ which is always the total demand $D_{s,t}$ of that scenario s. This simplification is helpful since it will reduce the calculation of the maximum regret if applicable.

To create a single robustness value $R(G_{i,t})$, the 90th percentile is taken P_{90} from the array $MR_{G_{i,t}}(s)$;

$$R(G_{i,t}) = \mathrm{P}_{90}\left(MR_{G_{i,t}}(s)\right), \quad s \in S \tag{4.12}$$

With $R(G_{i,t})$ the robustness value in ton H_2/day and $MR_{G_{i,t}}(s)$ the minimax regret of the network for each scenario.

This robustness value makes it possible to compare the network performance over different scenarios. The optimal value for $R(G_{i,t})$ is 0. This optimal value means that in 90% of the scenarios the network can facilitate the demand.

Example robustness measure

To illustrate the robustness measure a small example will be given by calculating the robustness for three network $G_{1,2055}$, $G_{2,2055}$, and $G_{3,2055}$ depicted in Figure 4.9. They are generated to fulfill the demand scenarios $D_{s,2055}$ for $s = 1, 2, 3$ depicted in Figure 4.10. It can be seen that scenario 3 is the only scenario where node 1 has demand, and this demand is high. As a result, the network $G_{3,2055}$ has high capacity pipelines towards node 1.

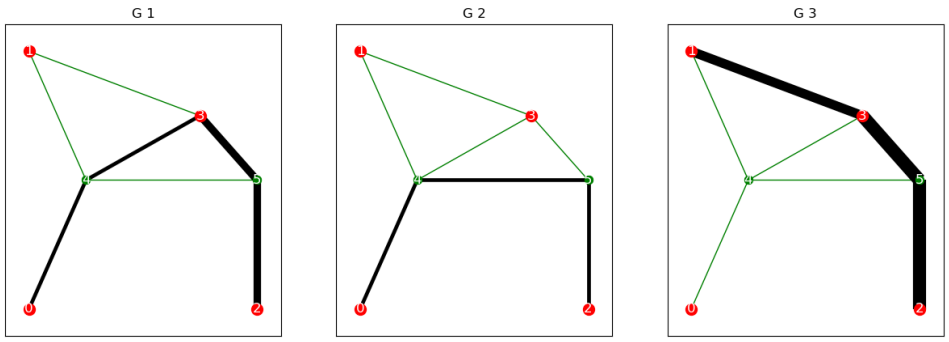


Figure 4.9: Networks 1,2, and 3 in 2055 used for the robustness measure example. The thickness of the edge indicates the capacity of the edge.

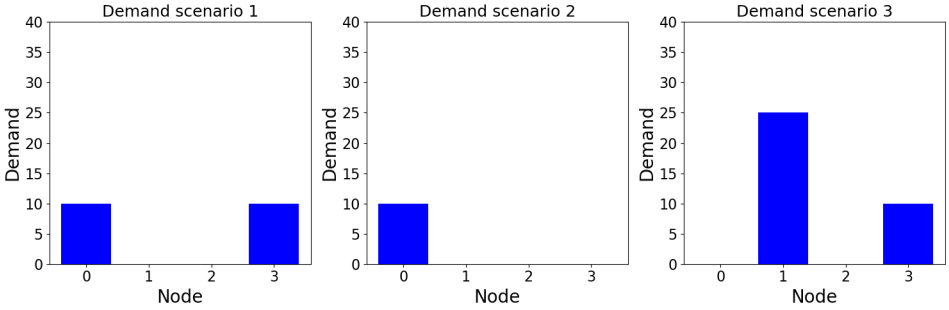
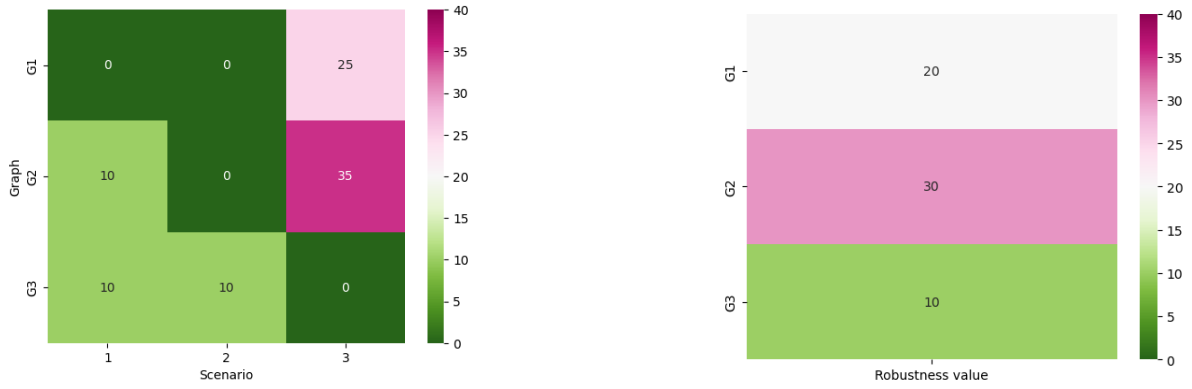


Figure 4.10: Demand scenarios 1,2, and 3 in 2055 used for the robustness measure example

The maximum regret $MR_{G_{i,t}(s)}$ is calculated and presented in an $i \times s$ matrix, as illustrated in Figure 4.11a. It is evident that the minimax regret is 0 when $i = s$, which is logical since this network is designed to meet that specific demand. Furthermore, it is observed that G_1 and G_2 perform poorly in scenario 3, as indicated by the last column in the heat map. Their underperformance can be attributed to the fact that they are not connected to node 1, which has high potential demand.



(a) Heatmap indicating the the minimax regret calculation of the three networks for three scenarios.

(b) Heatmap indicating the robustness value, the 90th percentile of the minimax regret for each network.

Figure 4.11: Example how the minimax regret and the robustness value are connected.

In this thesis, the focus is on decarbonization and facilitating demand; therefore, the minimax regret is considered in absolute terms rather than relative terms. This approach directly corresponds to the actual missed opportunities for hydrogen facilitation and indirectly influences the amount of CO_2 saved. As a result, the failure to facilitate a node with high demand (such as node 1) is weighted more heavily than the facilitation of a node with low demand (such as node 0).

Consequently, network G_3 emerges as the best-performing network, achieving a robustness value of 10, as illustrated in Figure 4.11b. Although network G_1 only fails to meet all demand in one scenario, this failure occurs in one-third of the total scenarios considered, which significantly impacts its robustness value. If G_1 were to fail to only once out of at least ten scenarios, its robustness score would be calculated as 0.

4.5. Robust Network Generation

In this section the robust networks heuristic is explained, these networks will later be used in the robust backtracking planning method (RBPM). The starting point of designing a robust network is that the robust network should have sufficient capacity to facilitate demand over time across multiple scenarios. Thereby once this robust network will be implemented the capacity will only be needed to be installed once, instead of every timestep.

For the generation of a robust network the scenario-optimized networks from the final timestep G_{i2055} are used as input. The rationale for selecting G_{i2055} as the input, rather than using the network directly calculated for the demand at $t = 2055$ without considering the path-dependency of previous networks, is that the detours made in the earlier time steps, while potentially sub-optimal for later periods, become beneficial when the robust network is implemented in the earlier time steps. These initial detours simplify decision-making by encouraging the use of shorter routes and reducing operational costs during the early stages.

The robust network will be developed in two steps, as illustrated in Figure 4.12. First, a robust topology is constructed from the scenario-optimized networks. Following this, capacity is assigned to each edge, resulting in the generation of a robust (capacitated) network. The methodology of initially selecting the most frequently occurring topology and subsequently assigning capacity to create a robust network is introduced by Huisman (2021). However, in this thesis, the facilitation of demand takes precedence over cost considerations, and the robustness metric is defined differently. Consequently, the process of assigning capacity is different. The details of these two steps will be explained in the subsections that follow.

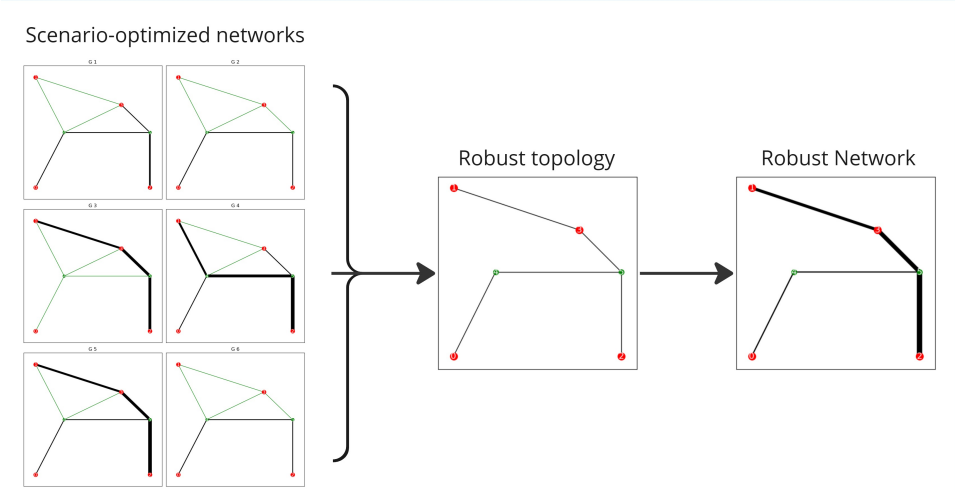


Figure 4.12: Schematic overview of the robust network generation steps

4.5.1. Robust topology

The robust topology is constructed by selecting all the edges present in the set of scenario-optimized networks $G_{i,2055}$, followed by the removal of cycles from this network. This is accomplished by eliminating the least frequently occurring edge and subsequently removing dead ends. A dead end is defined as an edge that terminates at a node, where that node is a Steiner node rather than an industrial node; in other words, the pipeline does not lead to any further connections.

To identify these dead ends, the algorithm iteratively examines each steiner node in the robust topology to determine if it has only one neighbor. If a Steiner node is found to have only one neighbor, the edge

connecting that node is removed.

The pseudo-code outlining this process can be found in Appendix D.1, while a visual example illustrating the robust topology heuristic is presented in Figure 4.13.

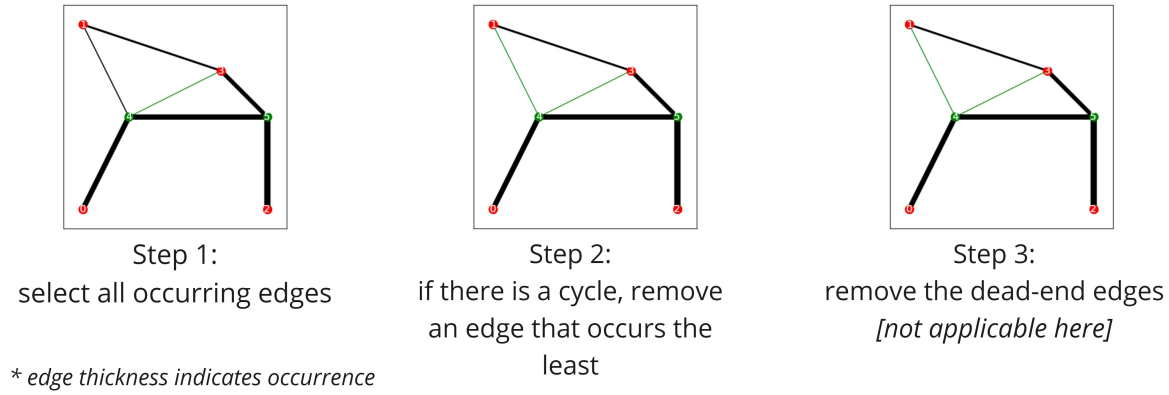


Figure 4.13: Visual example of the robust topology heuristic

4.5.2. Robust Capacitated Network

Given the topology, the robust network (RN) will be derived by assigning the robust topology capacity. As an important objective of the hydrogen network is that it should facilitate demand, the RN should score perfectly on the robustness measure ($R_{RN} = R_{RN,2055} = 0$). Since the robustness measure is a 90th percentile of the maximum regret this does not mean that the RN is equal to the maximal capacitated network. A maximal capacitated network is a network where the edge capacity is equal to the highest capacity it has been assigned in all the scenario-optimized networks $G_{i,2055}$. The maximal capacitated network is used to determine the RN by taking it as starting point and slowly lowering the capacity as long as the network does not raise the R_{RN}. This is done in the following steps:

Step 1 The maximum capacity network is derived by selecting for each edge e_k the maximum capacity of all $e_j \in G_{i,2055}(N, E)$ is selected. The maximum capacity network has the perfect robustness score $R_{RN} = 0$.

Step 2 For each edge e_k in the maximum capacity network the maximum capacity is lowered with increment capacity $\Delta_D Q$ as long as this decrease in capacity does not change the R_{RN}

4.6. Implementation Robust Network

The implementation of the robust network over time, referred to as the Robust Backtracking Planning Method (RBPM), will be described in this section. Backtracking, as introduced by André et al. (2014), serves as a strategic approach for deploying a projected network over time.

In the backtracking methodology, each timestep involves the installation of the necessary pipelines to meet the demand at various nodes, ensuring that the capacity aligns with that of the robust network. A significant advantage of planning in reverse, starting from the final robust network, is that it nearly eliminates the need to expand pipelines throughout the timesteps. In most scenarios, the robust capacity installed is sufficient to meet the projected demand in 2055. This approach contrasts with the immediate demand planning method, where the installed capacity is only adequate to address demand within that specific decade.

The pseudo code for the robust network implementation is presented in Appendix D.1.4. Figure 4.14 illustrates this implementation, demonstrating that no pipeline expansions occur, as evidenced by the absence of purple edges in the diagram.

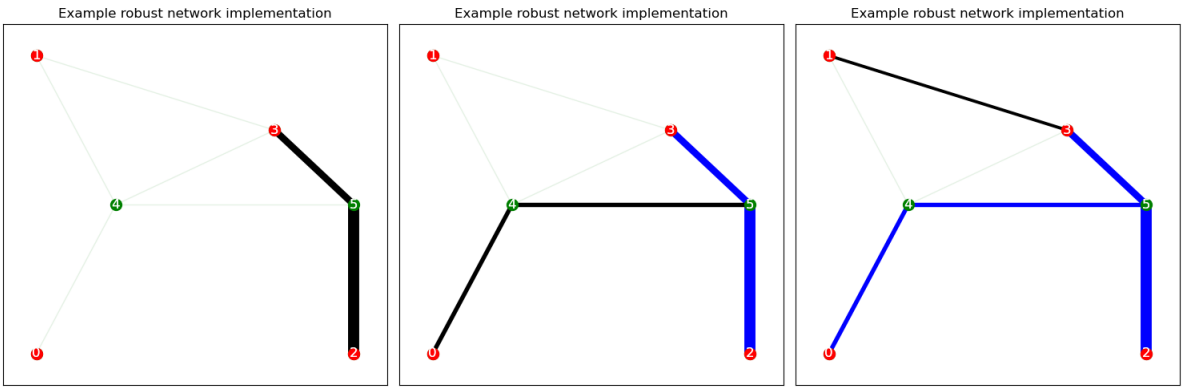


Figure 4.14: Example of the robust backtracking planning method, where the robust network is implemented for one future demand scenario

4.7. Comparing Planning Methods

The RBPM will be analyzed in comparison to two other planning methods: the immediate demand planning method (IDPM) and the minimum length robust backtracking planning method (ML-RBPM). Figure 4.15 illustrates the various planning methods and their implementations.

IDPM: The immediate demand planning method is a straightforward approach that involves installing only the pipelines necessary to meet demand for one decade. The IDPM matches the generation of the scenario-optimized networks in Section 4.3.

ML-RBPM: The minimum length robust backward tracking planning method (ML-RBPM) shares similarities with the RBPM. As indicated by its name, the robust topology is established by identifying the minimum length steiner tree within the graph and as result a minimum length robust network (ML-RN) is derived. The, reason why ML-RN is chosen as comparison is to investigate the difference between optimizing length as a topology and the RBPM, who indirectly optimizes costs over time by selecting the most occurring edges in the scenario-optimized networks.

Similarly to the RBPM method, the robust capacity is assigned in the same way. In Section 4.5, Step 1 is modified from selecting the emerging topology to deriving the MStT through the Kou heuristic (Kou et al., 1981). The kou heuristic does not differentiate between demand and supply nodes, resulting in a topology that is purely dependent on IPC layout rather than IPC configuration. In Appendix D.1, the pseudo-code for the ML-RBPM topology heuristic can be found.

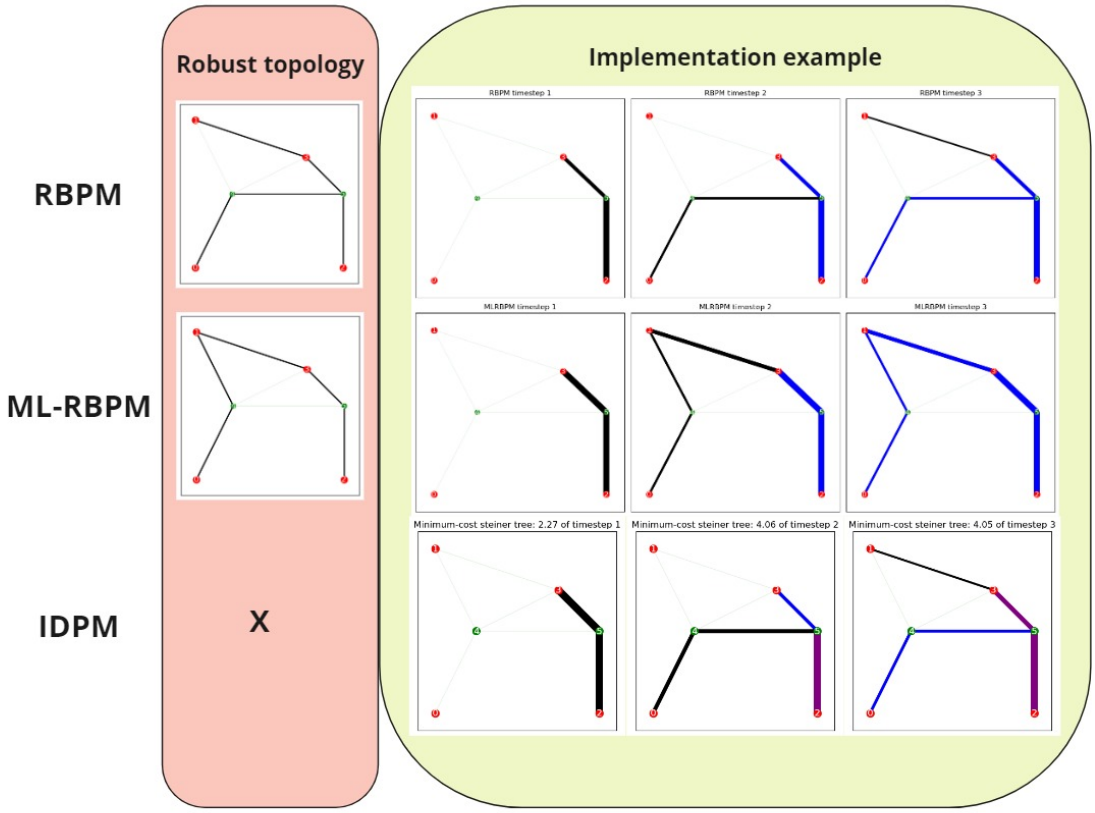


Figure 4.15: Overview of the three planning methods; Robust Backtracking Planning Method, Minimum Length Robust Backtracking Planning Method, and Immediate Demand Planning Method

5
Experimental setup

In this chapter, the data inputs for the method are described first, followed by the setup for each experiment in the second section, and finally, the performance metrics are presented in the last section.

5.1. Data inputs

The key data inputs described here are those needed to generate IPC configurations, as well as those used for optimization and analysis scenarios. An overview of the inputs and their role in the method can be seen in Figure 5.1. These inputs are described sequentially below.

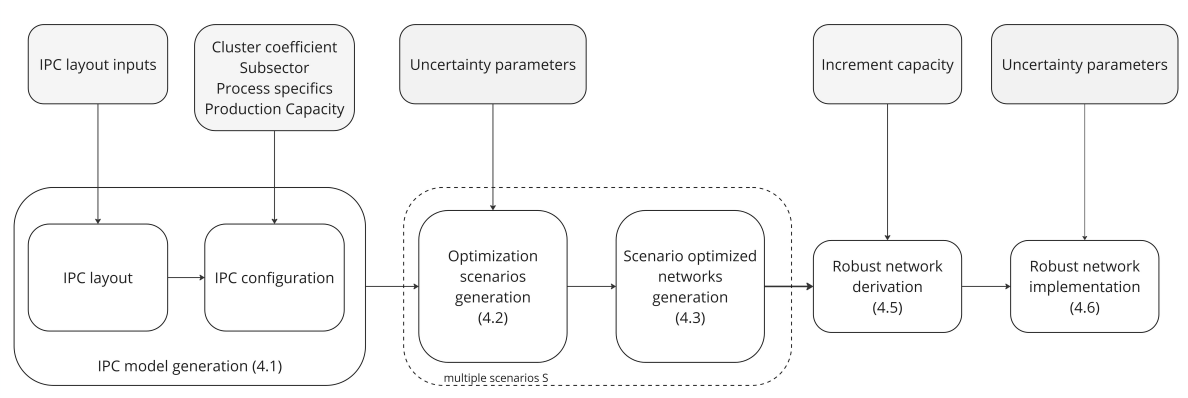


Figure 5.1: Overview of the inputs and where they are used in the method. The method steps are indicated by dotted squares and the corresponding section in the methodology is also given.

5.1.1. IPC layouts

An IPC layout consist of the location of the industrial nodes, a road network, and the wet infrastructure. For the analysis three IPC layouts are used; layout 0, 1, and 2. As described in Section 4.1, these layouts have been hand-selected to ensure that the wet infrastructure of the port visually aligns with existing ports. Table 5.1 provides an overview of the characteristics of each IPC layout. Although the three layouts contain an equal number of terminal nodes, they differ in size. Layouts 0 and 1 are relatively small, each covering approximately 6 km^2, which helps to limit computation time. In contrast, layout 2 has been designed to be larger in order to observe the behavior of the planning methods within a more extensive and complex road network. A visual representation of the three IPC layouts is depicted in Figure 5.2; however, it is important to note that the scale of layout 2 is larger than that of the other two layouts.

Table 5.1: Overview of the inputs for the IPC layout generation.

Parameter	Layout 0	Layout 1	Layout 2
Number of industrial plants	12	12	12
Number of existing demand nodes	9	9	9
Number of extra nodes	2	2	2
Number of supply nodes	1	1	1
Grid size road network	0.2 km	0.2 km	0.5 km
Maximum y coordinate supply node	0.8 km	0.8 km	2.5 km
Y dimension	3 km	3 km	10 km
X dimension	2 km	2 km	9 km
Area	6 km^2	6 km^2	90 km^2

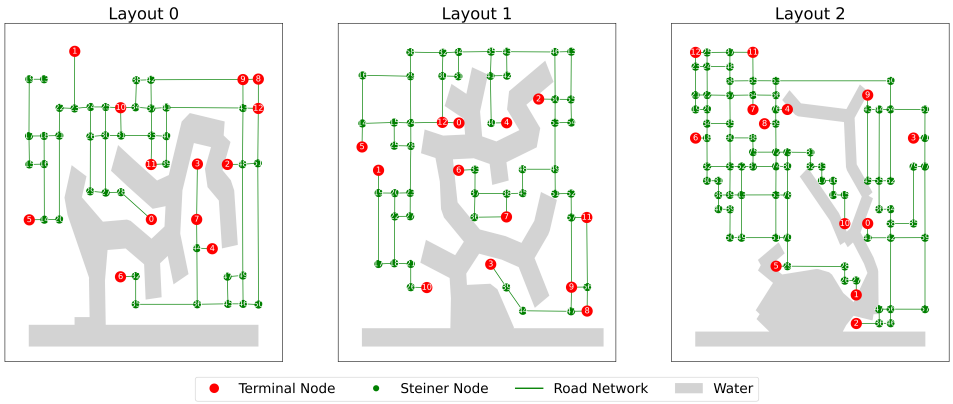


Figure 5.2: Overview of the three layouts used, green is the road network, red are the terminals and green are the steiner nodes. Layout 2 is around 10 times bigger than layout 0 and 1 and therefore the road network is more extensive.

5.1.2. IPC configurations inputs

An IPC configuration is generated by assigning plant and process specifics to the industrial nodes of the IPC layout, multiple configurations can be created for the same IPC layout. The IPC configuration is defined by the cluster coefficient, cluster sector, production capacity, and process characteristics. Below the values of these parameters are discussed.

Cluster coefficient

The cluster coefficient quantifies the proportion of specific subsector plants required to meet the cluster specification. In this thesis, the coefficient is a constant across all configurations, set at 0.5.This is assumption that at least 50% of the plants should be of the same (dominant) subsector is grounded in the desire for the IPC to have one dominant subsector that defines the specialization field of the industrial cluster as defined in section 2.1, while also maintaining heterogeneity in the IPC configuration.

Subsector

The subsectors are derived in chapter 2, encompassing hard-to-abate industries such as iron and steel, basic chemicals, and refineries, along with other sectors capable of implementing hydrogen-based heat technologies (Namazifard et al., 2024; Neuwirth et al., 2022b). A subsector is given to each terminal node. These subsectors are chosen randomly from the following list;

- Iron and Steel
- Basic chemicals
- Glass
- Metal processing
- Mineral processing
- Non-ferrous metals
- Pulp and paper
- Refineries

When generating an IPC configuration with a specific subsector as the specialization field, the industrial nodes are randomly assigned a subsector, and this process is iterated until 50% of the industrial demand nodes belong to the specific subsector.

Process Specifics

Process specifics are the attributes of industrial nodes assigned to demand nodes, including the process, temperature category, and prospective hydrogen decarbonization technology. Given the subsector the node belongs to a process from the relevant subsector is randomly selected, along with a hydrogen technology capable of decarbonizing that process. These selections are based on the specific energy consumption data created by Neuwirth et al. (2022a). An overview of this data can be found in Table 5.3. This data also includes the temperature range of the technologies, the technology readiness level, and the Specific Energy Coefficient (SEC), which are used in generating demand scenarios.

Production capacity

Next to a subsector, plants are also assigned a production capacity. The production capacity is derived from a process specific normal distribution. The parameters of the normal distribution for each process are presented in Table 5.2. This particular normal distribution is obtained from the German industrial plants database referenced in Neuwirth et al. (2022a), from which this thesis derives a normal distribution of the production capacity for each process of the existing German plants. A more detailed explanation of the derivation of production capacity is available in appendix E.

The normal distributions make it probable that a plant will be assigned a negative capacity. This is approached by establishing that if a randomly pick is negative, the minimal production capacity in the German database is selected.

Table 5.2: Production Capacity distribution as derived from the production data from Neuwirth et al. (2022a).

Process	μ [Mton/year]	σ [Mton/year]
Aluminum, primary	0.04	0.01
Ammonia	0.20	0.11
Board and packaging paper	0.02	0.02
Casting	0.17	0.12
Cement/Clinker	0.21	0.11
Chemical pulp	0.05	0.04
Chlorine, diaphragm	0.12	0.13
Chlorine, membrane	0.04	0.03
Container glass	0.02	0.01
Flat glass	0.04	0.02
Graphic paper	0.04	0.05
Lime burning	0.03	0.06
Methanol	0.08	0.03
Olefins	0.24	0.14
Refinery	0.85	0.55
Rolling (hot)	0.20	0.25
Steel, primary	0.51	0.26
Tissue paper	0.01	0.01

Table 5.3: Technology-specific energy consumption data (Neuwirth et al., 2022a)

Process	Technology	Temperature (°C)	Fuel SEC (MWh/t)	Feedstock SEC (MWh/t)	TRL (1–9)
Iron and steel					
Crude steel, primary	Blast furnace	1200–1450	3.2	-	9
	H_2-DRI	-	1.89	-	8
Metal processing					
Casting	NG burner	1100–1600	0.008	-	9
			0.028	-	9
	H_2-burner for process heat		0.028	-	4–5
Rolling (hot)	NG burner	700–1250	0.67	-	9
			0.58-0.61	-	9
			0.35	-	9
	H_2-burner for process heat		0.67	-	4–5
Non-ferrous metals					
Aluminum, primary	NG burner	700–950	2.1	-	9
			2.2	-	9
			Melting: 1.3	-	9
			Casting: 0.6	-	9
	H_2-burner for process heat		1.9	-	4–5
Glass					
Container glass	Recuperative NG burner	1450–1650	1.61	-	9
			1.28-1.72	-	9
	H_2-burner in furnace	1450–1650	1.28	-	4–5
Flat glass	Recuperative NG burner	1450–1650	3.03	-	9
			2.17-2.56	-	9
			2.58	-	9
	H_2-burner in furnace	1450–1650	2.17	-	4–5
Mineral processing					
Cement/Clinker	NG burner	1400–1450	0.97	-	9
			0.83-1.25	-	9
			1.08	-	9
	H_2-burner in rotary kiln	1400–1450	0.97	-	4–5
Lime burning	NG burner	900–1200	1.03	-	9
			1.14	-	9
	H_2-burner in furnace	900–1200	1.03	-	3–5
Pulp and paper					
Board & packaging paper	NG burner	80–220	1.36	-	9
			1.36-1.58	-	9
	H_2-burner in steam generator	80–220	1.36	-	8–9
Tissue paper	NG burner	80–220	1.92	-	9
			1.92-2.25	-	9
	H_2-burner in steam generator	80–220	1.92	-	8–9
Graphic paper	Ng burner	80–220	2	-	9
			2.0-2.33	-	9
	H_2-burner in steam generator	80–220	2	-	8–9
Recovered Fibers	NG burner	80–220	0.139	-	9
			0.139-0.167	-	9
	H_2-burner in steam generator	80–220	0.15	-	8–9
Chemical pulp	NG burner	130–150	3.42	-	9
			2.86-3.42	-	9
	H_2-burner for process heat	130–150	2.86	-	8–9
Basic chemicals					
Olefins	Steam cracker	800–950	6.64	-	9
			Boiler: 3.17-3.69	-	9
	Methanol-to-Olefins	-	17.67	-	8–9
Methanol	H2 from steam reforming	200–300	4.17	6.31	9
	H2 from electrolysis	-	6.31	-	8–9
Ammonia	Steam reforming	350–550	3.14	5.92	9
			2.5-4.61	-	9
	H_2 from electrolysis	-	5.92	-	9
Chlorine diaphragm	NG burner		0.86	-	9
			0.81-1.6	-	9
	H_2-burner for process heat		0.86	-	8–9
Chlorine membrane	NG burner for process heat		0.28–0.33	-	9
	H_2-burner for process heat		0.28	-	8–9
Refineries					
Refinery	H_2 for crude oil refining	-	0.389–0.639	0.389-0.639	9
	H_2 from electrolysis for hydro treating		-	-	7-9

5.1.3. Demand scenario generation

The data inputs discussed in this section are related to the demand scenario generation. One demand scenario consists of the demand of each industrial node for the years 2035, 2045, and 2055. For one configuration multiple demand scenarios are generated using the EMA work package (J. H. Kwakkel & Pruyt, 2013). The difference in these demand scenarios is created by the uncertainty parameters; demand range, transition probability, and extra node probability.

Other needed data inputs for demand scenario generation, such as SEC, TRL, and production capacity are set for every configuration and are therefore already known in this step. The demand generation function is tested to be deterministic for the same random seed.

Demand range

The demand range serves as a parameter for demand uncertainty, defining the extent to which the calculated demand may deviate from the actual demand (see equation 4.3). In the EMA work package, the demand range is represented as a categorical uncertainty parameter; for each generated scenario, one value from the predetermined set of demand range values is selected. This uncertainty parameter is set at 0.2 unless mentioned otherwise.

The value of 0.2 is established as an educated estimate. Given that the calculated hydrogen potential is derived from production capacity and specific process characteristics, along with the assumption that one process transitions completely at once, it is reasonable to conclude that the actual potential demand is likely to be situated close to the calculated demand. Consequently, a value of 0.2 has been chosen as a reasonable compromise that balances these considerations.

Transition probability parameters

The participant uncertainty parameters are the low, medium, and high transition probability parameters. All three are uncertainty parameters, assigned a range corresponding to the likeliness (low/mid/high) of future transition. The categorization of low, medium, and high is assigned according to the plant specifics, recall Section 4.2.2.

This categorization is first of all based on the temperature of the process. The data on technology-specific energy consumption provides the temperature range of a process, however, some technologies do not have a temperature indication. Table 5.4 illustrates that certain technologies, particularly in the basic chemical, iron and steel, and refineries sectors, lack an assigned temperature range since these technologies do not generate heat. Rather, there hydrogen is used as feedstock. Hydrogen is considered a viable option for the decarbonization of essential chemical, iron and steel, and refining processes (IEA, 2023a). Therefore, basic chemicals, refineries, and iron and steel are assigned a high transition probability.

Extra node probability parameter

Next to the transition probability parameters, there is also a extra node uncertainty parameter. This extra node probability is assigned to the terminal nodes that are labeled extra, and that dictates the chance of an extra plant appearing in the IPC. This extra node probability resembles the business climate, and the range of this probability is between zero and one, see 5.4.

Table 5.4: Overview participant uncertainty parameters range

Uncertainty parameter	Description	Minimum	Maximum
p_{lowT}	probability a process below **150** degrees transitions in one decade	0	0.125
p_{midT}	probability a process between **150 - 500** degrees transitions in one decade	0.125	0.3
p_{highT}	probability a process above **500** degrees transitions in one decade	0.3	0.5
p_{extra}	probabilty an extra plant appears on the grid in one decade	0	1

Participant uncertainty parameter range

Once the process is assigned a transition probability parameter, its range should be determined, as this range will be used by the EMA workbench to generate multiple scenarios. The uncertainty range, indicated by a minimum and maximum value, is shown in Table 5.4. The uncertainty parameter's range is reasonably assumed by examining the probability that a plant will transition to hydrogen by 2055, denoted as P_{2055}, as a function of the transition probability p_i of the process over a decade;

$$P_{2055} = 1 - (1 - p_i)^3. \tag{5.1}$$

After 30 years, we anticipate a high probability that high-temperature plants will have transitioned, while the likelihood of transition for low-temperature plants will be minimal. However, we do not desire the probability to approach 1, as this contradicts the inherently uncertain nature of the problem. Figure 5.3 illustrates the function of P_{2055} along with the selected maximum values for the temperature-based uncertainty parameters. The range is selected to ensure that the values are distributed with nearly uniform spacing across the P_{2055}.

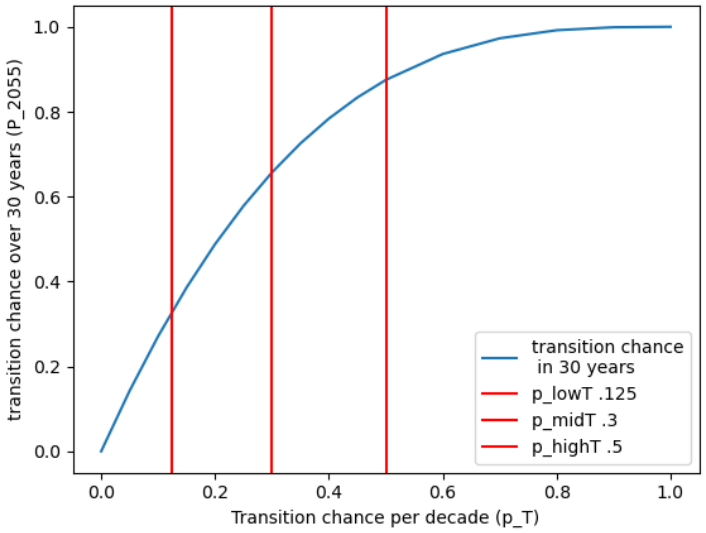


Figure 5.3: Indication of the maximum for each temperature-based uncertainty parameter and how that translates to the chance of the plants transitioning towards 2055.

5.1.4. Robust network generation input

Increment capacity

The increment capacity determines the accuracy regarding minimization of costs of the robust capacitated network, however it also heavily influences the computational time it takes to calculate the robust network. As a middle ground, the increment capacity is 0.01 ton H2/day. Compared with the total potential hydrogen demand in an IPC, this is between 5% and 0.5% of the total possible capacity of a cluster, depending on the dominant subsector in the cluster.

5.1.5. Robust network implementation

To evaluate the robust network implementation, new demand scenarios, referred to as analysis scenarios, are generated. It is crucial to emphasize that the optimization scenarios used to develop the network are distinct from the analysis scenarios. This distinction between optimization and analysis scenarios is significant, as the analysis scenarios should be viewed as test scenarios that do not replicate the training (optimization) scenarios.

Even when the ranges of uncertainty parameters are set identically, differences arise from the sampling methods employed by the EMA workbench, as well as variations in the random seed. These factors contribute to the divergence between optimization and analysis scenarios. Furthermore, it is possible to modify the uncertainty parameter ranges between analysis and optimization scenarios in experimental settings. This approach allows for the assessment of how the Robust Backtracking Planning Method (RBPM) performs in scenarios other than those for which they were optimized.

5.2. Set-up per experiment

In this section the experiments shall be described, including their link to the sub-research questions and their set-up. The experiment outcomes shall be presented in the next chapter, where in total one proof of concept and three experiments will be presented.

5.2.1. Proof of concept

How will this method be implemented for a real industrial port cluster?

First, a proof of concept to show how the method will be presented. Here it will be shown how the method can be applied to a real Industrial Port Cluster (IPC). Therefore, the proof of concept will apply the RBPM to one IPC configuration. Furthermore, this section will also give the reader a comprehension on the method applied to one IPC. This comprehension can be used as a little help to interpret the results of the other experiments which analyse a collection of multiple IPCs.

The chosen subsector for the proof of concept is Iron and Steel, this choice was made semi-arbitrary since it can be any subsector. The reason Iron and Steel was chosen is it has a lot of potential hydrogen demand, thereby it is easy to show the pipeline capacity changing over the years when more plants transition. The plant specifics of the configuration can be seen in Appendix H.1. The robust network is derived from 30 optimization scenarios, the robust network implementation with the RBPM is then illustrated with one analysis scenario.

Table 5.5: Set-up Proof of Concept

Parameter	Value
Method	RBPM
Layout	1
Cluster	Iron and Steel
Configuration	1
Optimization Scenarios	30
Analysis Scenarios	1

5.2.2. Experiment 1: Comparison simpler methods

How does the robust over time planning method perform compared to simpler planning methods?

The aim of this experiment is to answer the sub research question regarding the comparison of the robust plannings method (RBPM) with simpler methods. Therefore, in this experiment the performance of the RBPM is compared with the simpler methods; the minimum length robust backward planning method (ML-RBPM) and the Immediate Demand Planning Method (IDPM). The methods performance is measured in costs and length. Length is here chosen as a performance metric as length is used as optimization objective for the ML-RBPM robust topology heuristic, for more details recall Section 5.3. The three methods shall be evaluated for three layouts, as we want to test if the topology of the road network, and the placement of the terminal nodes affects the performance.

Furthermore, the experiment shall be conducted for two clusters; Iron and Steel, and Pulp and Paper. These clusters were selected due to the significant differences in their plant characteristics. Specifically, the Iron and Steel cluster exhibits a high demand for hydrogen and includes numerous processes with a high transition probability. In contrast, the Pulp and Paper cluster has a relatively low demand for hydrogen and predominantly consists of low-temperature processes characterized by low transition probabilities.

By analyzing these two contrasting clusters, the experiment aims to assess the performance of the methods while capturing a broad spectrum of plant specifics. The decision to limit the analysis to only two clusters is motivated by computational constraints and the desire to maintain clarity in visual representation and comparison. Including more clusters would not only increase computational complexity but also complicate the visualization of results. For each cluster, five configurations are tested, where each configuration has a different assignment of plant characteristics, the only thing constant is that at least 50% of the plant processes should belong to either Iron and Steel, or Pulp and Paper. An overview of the configurations used in this experiment can be found in Appendix H.2.

The robust network for the RBPM and ML-RBPM method are generated using 30 optimization scenarios, and the implementation analysis is conducted with 40 analysis scenarios.

Table 5.6: Set-up experiment 1

Parameter	Value
Methods	RBPM, IDPM, ML-RBPM
Layouts	0,1, 2
Clusters	Iron and Steel, Pulp and Paper
Configurations	5
Optimization Scenarios	30
Analysis Scenarios	40

5.2.3. Experiment 2: Performance over different clusters

What is the performance of the developed method across different clusters?

Goal of this experiment is to compare the performance of the robust backward planning method (RBPM) over different cluster subsectors and thereby looking for patterns of network development. This will be done for layout 0 and 1 for all different clusters (eight) and five plant configurations per subsector, in total 80 configurations. The reason why layout 2 is not analyzed is that the size of layout 2 reduces the computation time. A quick assessment of the configurations per cluster is provided below. A full overview of the plant specifics for layouts 0 and 1 can be found in Appendix H.3.

For each configuration a robust network is derived using 30 optimization scenarios. The RBPM performance is than analyzed with 40 different analysis scenarios, looking at costs. The demand range is kept constant at 0.2.

Table 5.7: Set-up experiment 2

Parameter	Value
Methods	RBPM
Layouts	0, 1
Clusters	Iron and Steel, Pulp and Paper
Configurations	5
Optimization Scenarios	30
Analysis Scenarios	40

Overview configurations

Overview of layout 1 and its configurations used for experiment 2 are depicted in two figures that show the plant characteristics of each configuration. Figure 5.4 displays the temperature category counts of the plants per configuration. As one can see, the distribution of temperature and feedstock processes over the plants differs per cluster, for some high probability processes have a high share (Glass, Basic Chemicals, Refineries and Iron and Steel) and for others medium probability processes are dominating (Non-ferrous metals), or low probability (Pulp and Paper).

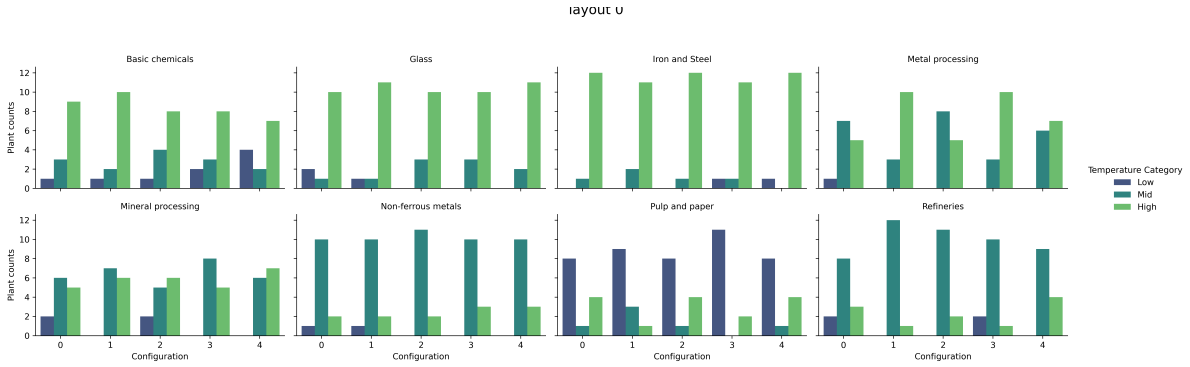


Figure 5.4: Overview of the amount of plants per temperature category for the 5 configurations per cluster for layout 1.

Additionally, Figure 5.5 displays the potential hydrogen demand per configuration, ordered per temperature category. The high probability processes dominate the potential hydrogen demand. This is because the average production capacity of certain high probability processes, such as Olefins, Refineries, and Steel, is approximately ten times higher than that of other processes (see Table 5.2). Consequently, these clusters exhibit significantly high potential hydrogen demand, while Glass, despite having a substantial share in high probability processes, has lower total potential demand.

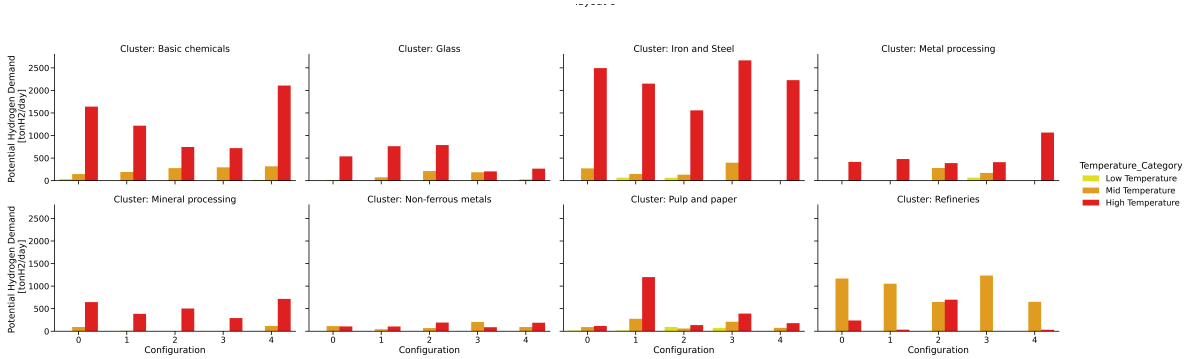


Figure 5.5: Overview of the total potential demand per cluster, with on the x-axis the 5 configurations for layout 1 per cluster.

5.2.4. Experiment 3: Performance under higher demand uncertainty

How does the developed method perform under different demand uncertainty ranges?

Goal of this last experiment is to access the performance of the RBPM under a higher uncertainty of demand. Until now, it is assumed that the demand range parameter is a constant (0.2); meaning there is some uncertainty in the asked demand per node over the scenarios, namely ± 0.2. However, it is possible that this range in demand uncertainty is larger or smaller. Furthermore, it is possible that the robust network will be trained with optimization scenarios that are set to a certain demand range while in reality another demand range is the case. Therefore, in this experiment different demand ranges will be tested for both optimization scenarios and analysis scenarios.

The demand range is varied for both the optimization and analysis scenarios separately. Therefore, the optimization demand range and the analysis demand range analysis is introduced. The optimization demand range optimization is a constant that is set per robust network optimization, the analysis demand range is a categorical uncertainty parameter, where using the EMA worckpackage for each scenario one value from the possible categories is chosen as that demand uncertainty. This way, the robust network optimized for one demand range can be tested for a variety of demand range values in the analysis scenarios.

We distinguish two sub-experiments. Experiment 3a looks at what happens if the demand uncertainty is higher in reality than optimized for. Experiment 3b studies what happens if the optimized demand

uncertainty is high and the real demand uncertainty is as much or less. The set-up for both experiments is listed below.

Table 5.8: Set-up experiment 3a

Parameter	Value
Methods	RBPM, IDPM
Layout	1
Cluster	Iron and Steel
Configurations	1
Optimization Scenarios	60
Analysis Scenarios	90
demand range optimization	0.2
demand range analysis	[0.1,0.2,0.3,0.4,0.5]

Table 5.9: Set-up experiment 3b

Parameter	Value
Methods	RBPM, IDPM
Layout	1
Cluster	Iron and Steel
Configurations	1
Optimization Scenarios	60
Analysis Scenarios	90
demand range optimization	0.5
demand range analysis	[0.1,0.2,0.3,0.4,0.5]

The highest demand range setting in these experiments is 0.5, this value is already quite high, but chosen for exploratory reasons; to see what happens in the extreme. The reason why not an even higher value because this would increase the amount of optimization scenarios too much; thereby increasing the computational time too heavily. Additionally, if a demand range of more than 50% would be practiced, in reality this would mean that a plant would increase or decrease its production capacity with more than 50%, which is an unlikely scenario.

Furthermore, the RBPM performance is compared with the IDPM performance to access more clearly the response of the RBPM. Mainly, because it is interesting to see the difference between the RBPM that is made to be robust for all scenarios versus the IDPM method that is fit for only one specific scenario. The RBPM and IDPM will be compared regarding total costs, costs per decade, demand facilitated and demand not facilitated.

5.2.5. IPC configurations per experiment

The configuration specifications of the setups for each experiment are presented in Table 5.10, where the total configurations equal the product of the number of layouts, the number of clusters, and the number of configurations per cluster.

Table 5.10: Overview amount of configurations per experiment

	Layouts	#clusters	cluster specification	#configurations per cluster	total configurations
Proof of concept	1	1	Iron and Steel	1	1
Experiment 1	0,1,2	2	- Iron and Steel - Paper and pulp	5	30
Experiment 2	2	8	- Iron and Steel - Paper and pulp - Glass - Refineries - Basic chemicals - Non-ferrous metals - Metal processing - Mineral processing	5	80
Experiment 3	1	1	Iron and Steel	1	1

5.3. Performance metrics

The performance metrics used to measure and compare the methods are presented in this section. The metrics relate to costs, demand facilitation, and duration. These performance metrics overlap with the objectives of cost and demand facilitation used for the developed method, which is logical, as they should align with the actors' objectives derived in chapter 2. Nonetheless, all performance metrics are mentioned below for completeness, and any (dis)similarities with the objectives will be noted.

5.3.1. Costs

Costs are an important objective for the actors, see chapter 2. In the IDPM, costs are minimized for every decade. The RBPM and ML-RBPM do not have a direct cost objective, however, by installing more capacity than needed in the first decades, it is expected that the (ML-)RBPM still reduces costs. To grasp the different behaviour of these planning methods both costs per decade (C) and cumulative costs (CC) are used as performance metrics. Additionally, to compare the methods better the relative cost function (RC) is also introduced.

Costs per decade

The function costs per decade C is similar to the cost function described in the methodology;

$$C^{m}_{l,c,s}(t) = ((1 + (0.04 \cdot 5) \cdot C^{C,m}_{l,c,s}(t) + 0.04 \cdot \sum_{i=2035}^{t-10} C^{C,m}_{l,c,s}(i) \tag{5.2}$$

where $C^{m}_{l,c,s}(t)$ is the costs of decade t in M euro for layout l, configuration c, and scenario s when using method m. 0.04 indicate the fraction of operational costs, and $C^{C,m}_{l,c,s}(t)$ is the capital costs for the installation of the network using method m for layout l, configuration c, and scenario s at timestep t, the capital costs are calculated with equation 4.7

Cumulative costs

The cumulative costs CC provide a clearer overview of the total incurred over multiple decades compared to the costs calculated per decade. This perspective is particularly useful for analyzing the total costs made by the year 2055 for each method.

$$CC^{m}_{l,c,s}(t) = \sum_{i=2035}^{i=t} C^{m}_{l,c,s}(i), \text{ with } t = 2035, 2045, 2055 \tag{5.3}$$

where the cumulative costs $CC_{l,c,s}$ in M euro is the summation of costs per decade, indicating the total costs spend on the network generated by method m up until decade t. The cumulative costs are calculated per layout l, configuration c and scenario s.

Relative costs

The relative costs RC are particularly useful when comparing different methods with each other. The relative costs is the difference of the cumulative costs of one method m with the $RBPM$. A positive value of RC means that the RBPM has lower cumulative costs compared to method m.

$$RC^{m}_{l,c,s}(t) = CC^{m}_{l,c,s}(t) - CC^{\text{RBPM}}_{l,c,s}(t) \tag{5.4}$$

where $RC^{m}_{l,c,s}(t)$ is the difference in cumulative costs in million euros between method m and the RBPM at time step t.

5.3.2. Facilitation of demand

Demand facilitated is next to costs the other objective derived from the actor analysis in chapter 2. For performance measurements, both the demand facilitated is measured, as well as the demand *not* facilitated.

Demand facilitated

The facilitated demand is defined in the same manner as the methodology. The equation is:

$$FD^{m}_{l,c,s}(t) = \Phi(N^{m}_{l,c,s}(t), D_{l,c,s}(t)) \tag{5.5}$$

Where $FD^{m}_{l,c,s}$ is the facilitated demand in ton H_2/day at timestep t, and $\Phi(N^{m}_{l,c,s}(t), D_{l,c,s}(t))$ the maximum flow the network $N^{m}_{l,c,s}(t)$ can facilitate for that specific demand scenario $D_{l,c,s}(t)$ in ton H_2/kg.

Demand not facilitated

In addition to facilitated demand, we are also interested in non-facilitated demand NFD. Examining non-facilitated demand provides a comprehensive overview of missed opportunities related to hydrogen that are not currently being facilitated. Additionally, it gives a quick comparison between the total demand in a scenario and the facilitated demand. Non-facilitated demand is defined as

$$NFD^{m}_{l,c,s}(t) = D_{l,c,s}(t) - FD^{m}_{l,c,s}(t) \tag{5.6}$$

Where $NFD^{m}_{l,c,s}$ is the not facilitated demand in ton H_2/day at timestep t using method m applied to layout l, configuration , $D_{s,t}$ is the demand for scenario s at timestep t, and $\Phi_{G^{m}_{l,c,s,t},D_{s,t}}$ the maximum flow the network can facilitate for that specific demand scenario $D_{l,c,s}(t)$ in ton H_2/kg.

5.3.3. Length

Length is used as a performance metric, despite it not being an objective of the actors. This metric is still used as the difference between the topology heuristics RBPM and ML-RBPM is that RBPM has a occurrence heuristic and ML-RBPM a minimum length heuristic. Therefore, when comparing these methods, the length is also chosen as a performance metric, to assess the difference. The length metric is defined as follows;

$$L^{m}_{l,c,s}(t) = \sum_{(i,j)\in N^{m}_{l,c,s}(t)} l_{ij} \tag{5.7}$$

6 Results

In this chapter, the results are presented as follows: First, the proof of concept for the developed method is introduced. Next, the performance of the Robust Backtracking Planning Method (RBPM) is compared with the Immediate Demand Planning Method (IDPM) and the minimum length robust backtracking planning method (ML-RBPM) in Experiment 1. This is followed by Experiment 2, which compares RBPM across different clusters. Lastly, Experiment 3 evaluates the performance of the RBPM under various types of demand uncertainty.

6.1. Proof of concept

How will this method be implemented for a real industrial port cluster?

The configuration used as proof of concept is an Iron and Steel cluster with layout 1. The characteristics of each node for this specific configuration are presented in Table H.1. First, a visual analysis of the optimization scenarios is presented, and then the robust network and its implementation is presented.

6.1.1. Visual analysis

The 30 scenario-optimized networks are visualized in a capacity graph (Figure 6.1) and an occurrence graph (Figure 6.2). The overlapping capacity graph illustrates that pipeline capacity expands over time. The thickness of an edge indicates the capacity through the pipes, and the blackness of the edge indicates the frequency that capacity is set to the edge.

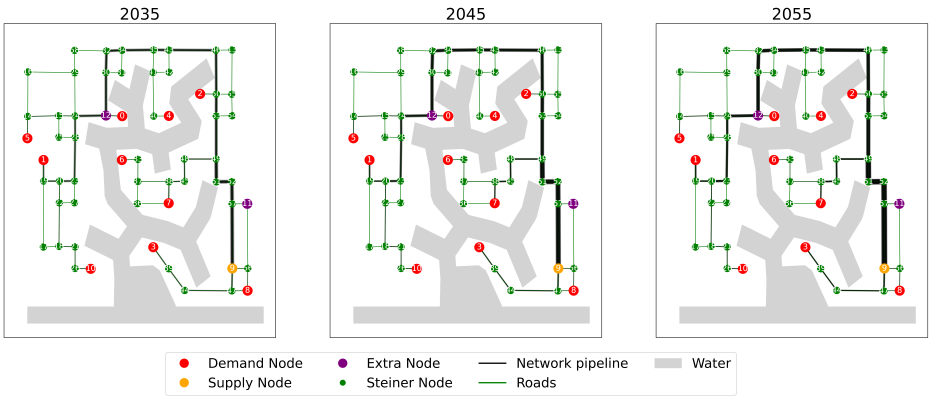


Figure 6.1: Overlapping capacity graph of all the 30 demand networks. The width of the edge indicates the capacity and the darkness of the edge indicates how often the edge in that capacity is used.

The occurrence graph illustrates the frequency of each edge, with the thickness of the edges representing their relative frequency. The graph indicates the emergence of a network tree, characterized by thicker edges. However, a small cycle is visible in the bottom-right corner of the occurrence graph. This cycle arises from the optimization of the network for each demand scenario timestep, where the connections between nodes vary depending on whether they participate in the scenario simultaneously.

Since the costs of installing a new pipeline are higher than those of installing a pipeline with greater capacity, the most cost-efficient network in 2055 will favor the sharing of pipelines. This principle is illustrated in the utilization of the left and lower sides of the graph. Notably, the lower pipeline in the graph appears the least frequently, which can be attributed to the asynchronous transitioning of the plants.

Once a pipeline is installed with a specific capacity, any additional capacity is treated such that it necessitates the installation of new pipelines. Consequently, when the nodes transition asynchronously, they are regarded as two separate optimization problems. This results in a different network configuration that minimizes the distance between the nodes and the supply node compared to scenarios in which transitions occur synchronously. In the latter case, the IDPM algorithm would perceive the situation as a combined optimization problem, where shared capacity could lead to a more significant reduction in costs compared to a small decrease in pipeline length.

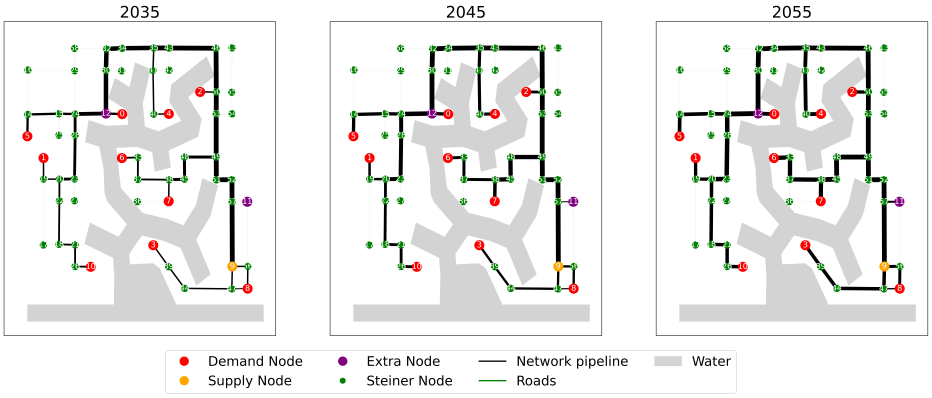


Figure 6.2: Occurrence graph visualizing the occurrence of the edges by the thickness of the edge.

6.1.2. Robust network

The robust network, constructed from scenario-optimized networks described is illustrated in Figure 6.3. The robust network looks the same as the most occurring edges in the occurrence graph. It can be seen that the cycle in the bottom right corner is removed; however, the most cost-efficient trajectory, at the bottom, of the cycle is not selected. The advantage of this choice is that if only one plant transitions in a plausible future, the costs are lower; however, if both plants transition, the costs are suboptimal.

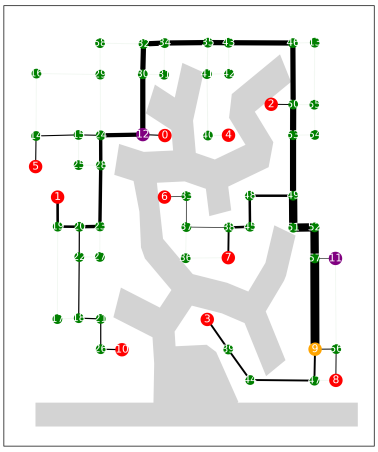


Figure 6.3: Overview of the robust network derived by applying the robust algorithm to the case example

In Figure 6.4, one can see how the robust network is implemented for one possible future scenario. It can be seen that once pipelines are installed they are not extended over the years, this is an advantage as this will limit installation costs. This is in contract with the immediate demand planning method (IDPM), where often pipelines are extended. In Figure 6.5the IDPM implementation is shown for the same future scenario, the purple pipelines indicate the extended pipelines.

Furthermore, looking at the costs per method for one scenario in Table 6.1, one can see the difference between the IDPM and the (ML-)RBPM. Where RBPM have relatively high costs in 2035, but less costs in 2045 and 2055 compared to the IDPM. This behaviour will be further studies in the next experiment.

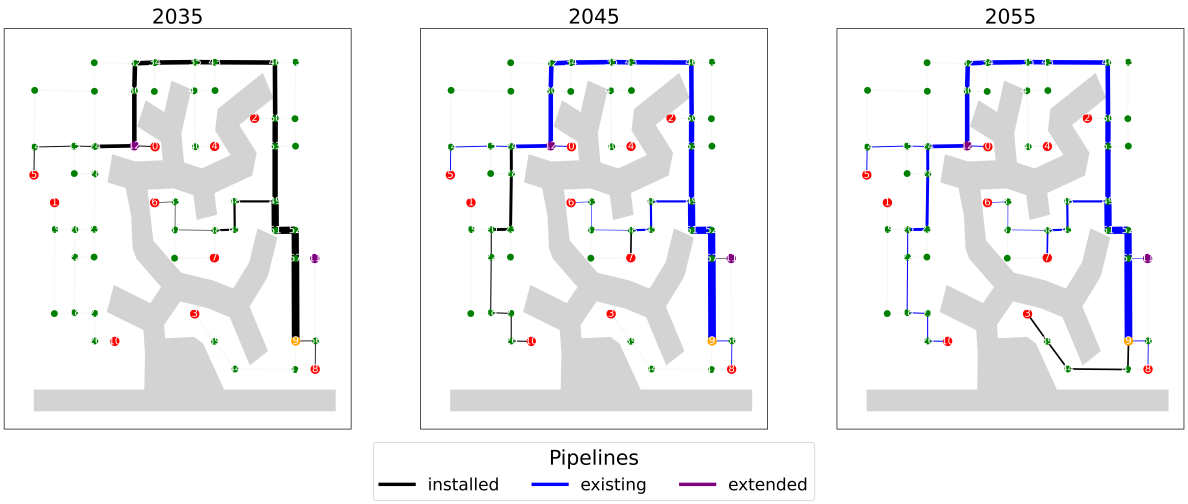


Figure 6.4: Development of a network in a hypothetical future using the Robust Backtracking Planning Method.

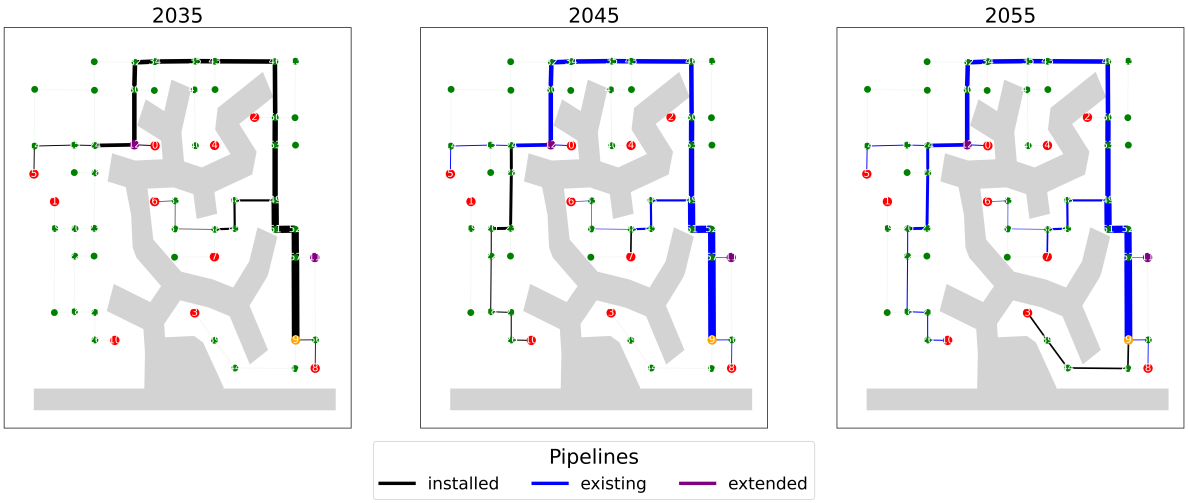


Figure 6.5: Development of a network in a hypothetical future using the Immediate Demand Planning Method.

Table 6.1: Costs per decade and total costs for implementing network for one analysis scenario for the three different methods.

Method	2035	2045	2055	Total
IDPM	120	130	120	370
ML-RBPM	160	80	80	320
RBPM	150	80	80	310

6.2. Experiment 1: Comparison simpler methods

How does the robust over time planning method perform compared to simpler planning methods?

Here, the RBPM is compared with the two simpler methods ML-RBPM, and IDPM, who are defined in section 4.7. They are compared by looking at both Iron and Steel and Pulp and Paper clusters for three different layouts. Layout 0 and 1 represent both a relatively small IPC area (6km^2), and layout 2 represents a large IPC area of around 90 km^2. The methods are compared by looking at costs, and length. First the length comparison shall be presented, and secondly the cost comparison.

6.2.1. Length

The three methods are compared by looking at the length in 2055 of the different networks implementation, per layout and for the five configurations per layout. The behaviour is the same for each layout, but for layout 2 the pattern is enlarged due to the bigger area covered by the IPC. In Figure 6.6 the length comparison is shown for layout 2 per configuration and per cluster, the length comparison for all layouts can be seen in appendix Figure I.1. Notably, the RBPM scores better on length for every layout, compared to the ML-RBPM. This is a surprise since the robust topology heuristic of ML-RBPM, the MStT heuristic, minimizes length, while RBPM topology heuristic maximizes the occurrence of edges from the optimization scenarios.

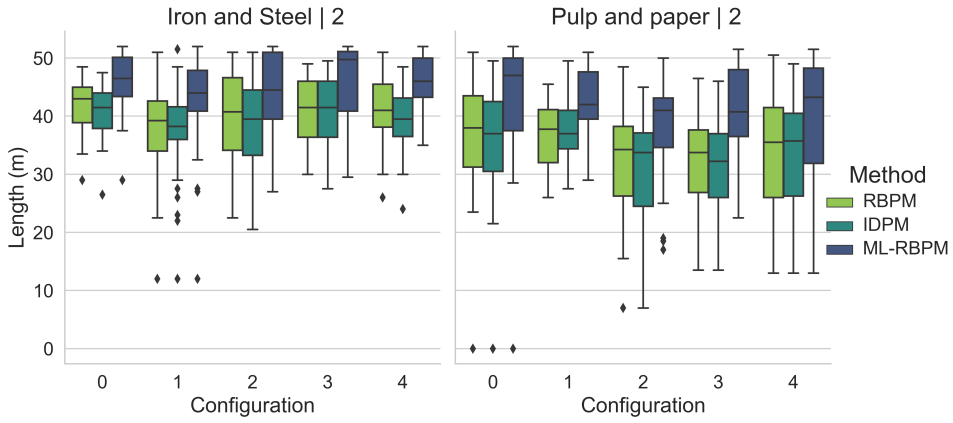


Figure 6.6: Boxplot comparing the lengths for the networks developed by either RBPM, IDPM, or ML-RBPM.

In Figure 6.7, the robust network (RN) topology and the minimum length robust network (ML-RN) topology are visually shown for a configuration 3 with supply node 2. The length of the ML-RN is smaller (52km) than the RN who is 53.5km. So here it can be seen that the heuristic has minimized length for the ML-RN compared to the RN. However, when the robust networks are deployed over time using backtracking the length differences are created.

Namely, when an extra node does not participate in a future scenario, this has effect on the length of the network. Specifically on the RBPM network as there the extra nodes are connected to the network with branches, meaning that these branches only need to be placed when the extra nodes participate. However, this is not the case for ML-RBPM, since the extra nodes are part of trajectories connecting the supply node with other terminals, it does not matter whether the extra node participates, the pipelines will most likely be placed.

Therefore, when executing the ML-RBPM and RBPM, the networks deployed with RBPM have on average a lower total length than the ML-RBPM. Zooming in on the outliers of the RBPM, in Figure 6.6, one can see that these outliers are almost equal to the highest values of the ML-RBPM. Since these lengths are the lengths when almost all extra nodes are participating in the network.

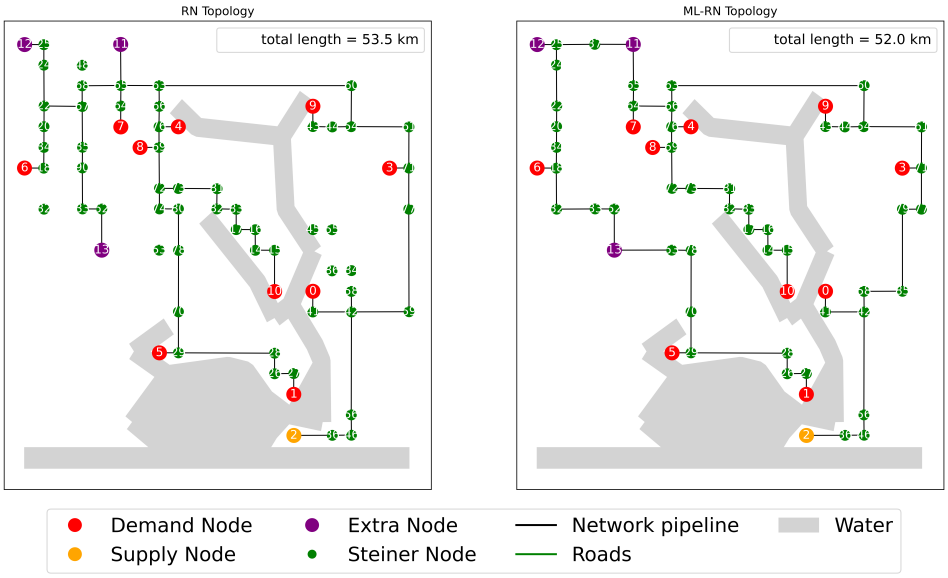


Figure 6.7: Comparison of the robust topology of the robust network (RN) topology used for RBPM and the minimum length robust network (ML-RN) topology used for ML-RBPM. This comparison is made for the robust network developed for Iron and Steel cluster, configuration 3.

6.2.2. Costs

To assess the difference between the simpler methods with the RBPM method regarding costs, the relative costs RC are presented in Figure 6.8 per configuration. A positive value of RC means that the RBPM performs better regarding costs in that decade. The cumulative costs comparison can be found in appendix I.1.1. Below we will first discuss the behaviour of ML-RBPM, and then discuss the IDPM compared to the RBPM.

Cost comparison with ML-RBPM

In Figure 6.8 the $RC^{ML-RBPM}$, indicated as the dark blue lines, is always positive, and therefore performs worse than the RBPM regarding costs. The cause of this is already explained in the section above; the ML-RBPM topology does not distinguish supply, extra, and demand nodes and therefore the ML-RN is not tailored for the specific configuration. Whereas, the RBPM selects the most occurring edges of the optimization scenarios and thereby the selected topology is tailored for the specific configuration.

Zooming in on layout 2, one can see two patterns in the $RC^{ML-RBPM}$; for configurations 0, and 3 the relative costs are maximum 150 million euro, while the $RC^{ML-RBPM}$ for configuration 1,2 and 4 are around 200 million euros or higher. This difference is caused by the different placement of the supply node between these patterns. Both configuration 0, and 3 have supply node 2, corresponding with the supply node on the right of the water, and the other configurations have the supply node on the left side of the water, see Figure 6.7.

In appendix I.1.2, the RN and ML-RN of the five configurations are presented. And from there it is observed that for configuration 0, and 3 the RN and ML-RBPM topologies have a more similar pipeline structure and specifically the pipeline structure with high capacity. Thus for these configurations the costs of RBPM and ML-RBPM are more alike and therefore the $RC^{ML-RBPM}$ is more closely to zero. However, the RBPM still performs better for every layout, cluster, configuration and at any timestep.

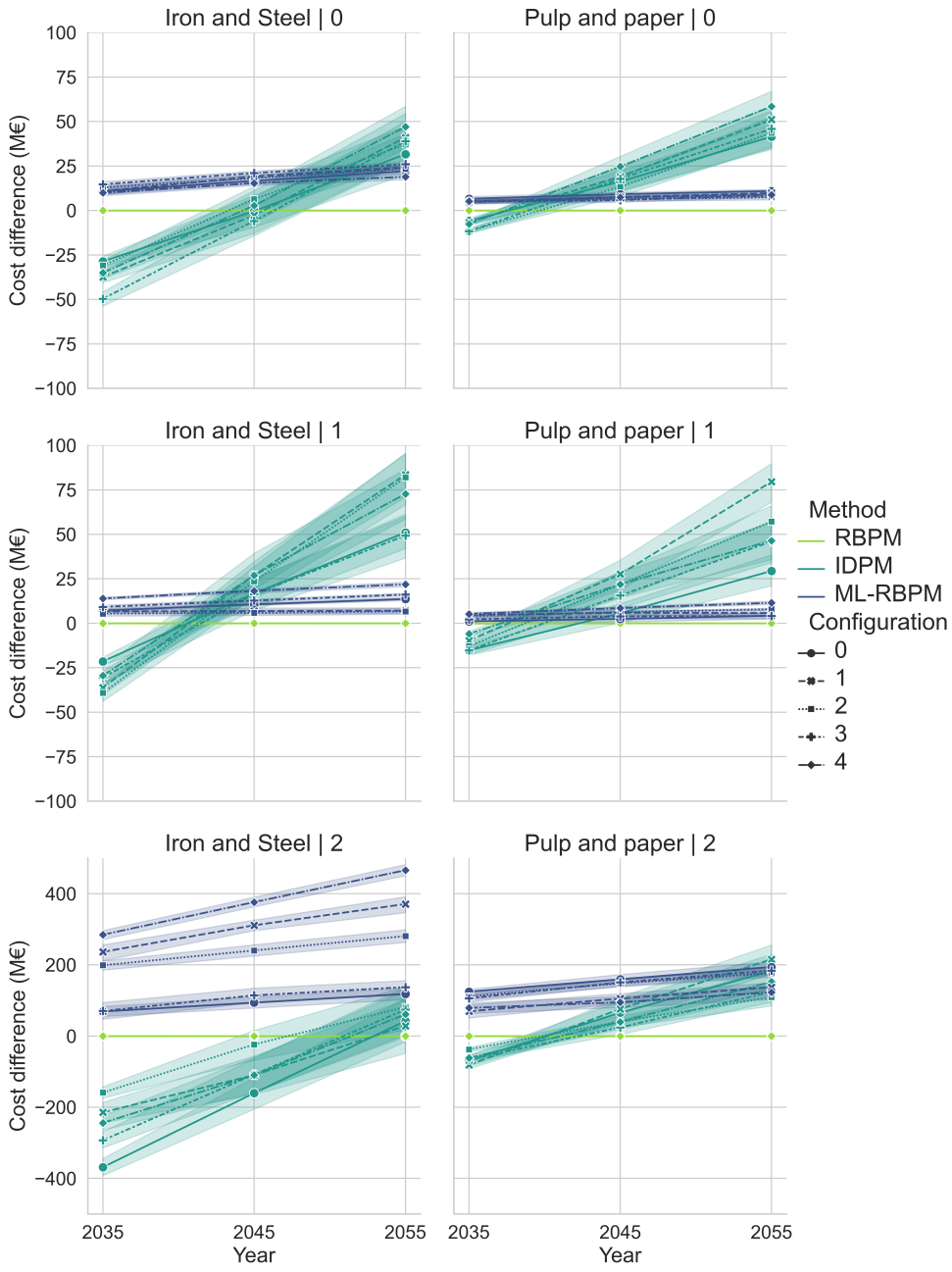


Figure 6.8: Relative cost difference comparing the cumulative costs of the methods with the cumulative costs of RBPM per decade. A negative value for relative costs, means that in that decade the method has lower costs than RBPM.

Cost comparison with IDPM

After the cost comparison of the ML-RBPM, the IDPM comparison is presented in this section. First thing that stands out is that the $RC^{IDPM}(2035)$ is negative for every cluster, layout configuration. The cost for the IDPM network in 2035 are lower compared to the RBPM costs. This is caused by the fact that the pipelines placed in 2035 for the RBPM have an overcapacity, while the IDPM placed pipelines are just the right fit for the asked demand at that timestep. Every decade new pipelines need to be installed with IDPM to extend the capacity, while the costs of RBPM after 2035 are less since no pipelines need to be extended and thus there are less installation costs. As a result the RC^{IDPM} grows. It differs per layout, and cluster, at which timestep the RC^{IDPM} is positive. But in almost all cases $RC^{IDPM}(2055)$ is positive, indicating that RBPM results in less total costs spends over 30 years, than IDPM. Below the differences in behaviour of RC^{IDPM} for different clusters and layouts is discussed

Looking at the difference in clusters one can already see that the RC^{IDPM} is steeper in the Iron and Steel clusters compared to the Pulp and Paper clusters. Secondly, one can see that the starting point $RC^{IDPM}(2035)$ is lower for Iron and Steel compared to Pulp and Paper. This different pattern is caused by the difference in hydrogen potential demand in the clusters. Pulp and paper is the sector with the least hydrogen potential demand, while Iron and Steel has the most, see section 5.2.3. Therefore in an Iron and Steel cluster, the initial costs in 2035 for RBPM are very high as much capacity is installed at once. While in the Pulp and Paper sector the extra capacity that is installed is not that much which explains the difference in $RC^{IDPM}(2035)$. The steepness is also caused by the higher potential demand, as the extended pipeline capacity required for IDPM is greater, leading to a more significant increase in CC^{IDPM} over time compared to CC^{RBPM}.

Additionally, it is interesting to note that by 2055, the total costs saved with the implementation of RBPM for Pulp and Paper are comparable to those for Iron and Steel. Although the absolute cumulative costs for Iron and Steel are higher, see appendix Figure I.2.

Looking at the RC^{IDPM} over the different layouts for the same clusters, so comparing vertically in the Figure 6.8, the layout 2 for Iron and Steel shows much more diversion for RC^{IDPM} over the configurations, and additionally the $RC^{IDPM}(2055)$ is near zero. In contrast, layout 2 for Pulp and Paper shows exactly the same pattern as layout 0, and 1. Per layout, some slight difference in the spread of configurations can be seen, for Pulp and Paper this is likely caused by the specific plant specifics of these configurations

6.3. Experiment 2: Performance over different clusters

What is the performance of the developed method across different clusters?

In this experiment the RBPM performance is analyzed for different clusters. For this analysis the main focus is on seeing patterns in the RBPM cost behaviour and linking this to the cluster specifications. These clusters and the configuration characteristics are depicted in Appendix H.3.

To look at the difference the costs per decade $C(t)$ are compared over different clusters for RBPM. In Figure 6.9, a box plot is provided for a comprehensive overview of the cost per decade, illustrating the distribution of values for each cluster, combining data from all layouts and configurations. It can be seen that Iron and Steel, Refineries, and Basic chemicals have the highest costs in 2035 $C(2035)$. The other clusters show less initial costs in 2035, and in comparison also less or equal costs in 2045 and 2055.

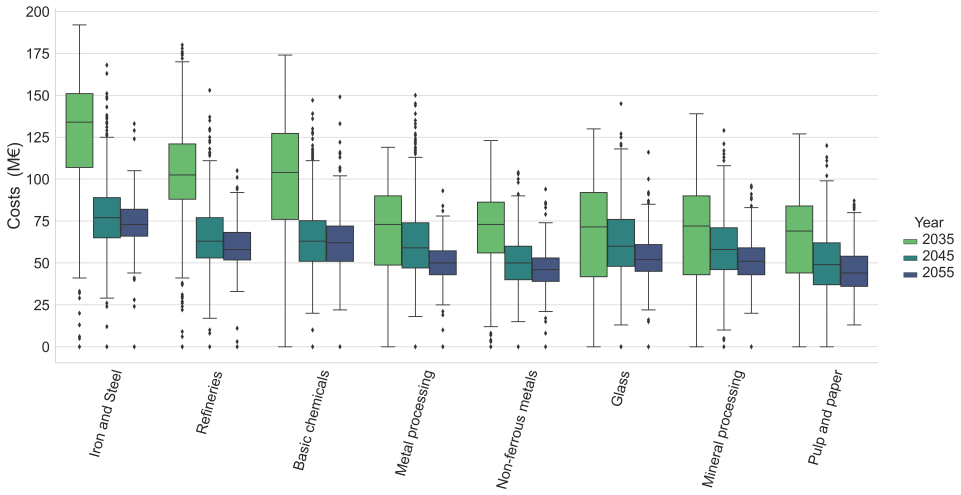


Figure 6.9: Costs per decade for the implementation of RBPM, the presented data is a collection of the data from layout 0 and 1.

Furthermore, a pattern for each cluster can be seen where $C(2045)$ and $C(2055)$ show less variation in costs, compared to the high variation in $C(2035)$. This is as expected, as the more pipelines are already placed of the robust network to facilitate certain nodes, there is less variation possible towards 2055 for placing extra nodes and thereby generating extra costs.

As noted in the previous experiment, configuration specifics significantly impact costs. Figure 6.10 illustrates the costs per decade for layout 1, categorized by cluster. The costs vary among configurations within each cluster, particularly for mineral and metal processing, where differences are pronounced. For instance, in metal processing, Configuration 1 shows that over 25% of scenarios have zero costs projected for 2035.

Furthermore, in Appendix I.2, the costs per decade for layout 0 are presented, revealing the same patterns observed in layout 1.

Figure 6.10: Cost per decade RBPM for layout 1. For each cluster, the costs per decade are depicted for each configuration

6.4. Experiment 3: Performance under higher demand uncertainty

How does the developed method perform under different demand uncertainty ranges?

In this experiment, the behaviour of the RBPM under higher uncertainty is tested by changing the demand range. First, in experiment 3a the RN is not optimized for this increased demand range, and then, in experiment 3b the RN is optimized with this increased demand range.

6.4.1. Analysis scenarios with higher uncertainty

In Figure 6.11, the kernel density estimation plot of the costs and facilitated demand can be seen for both the IDPM and the RBPM for all demand range scenarios. The RBPM is trained with demand range Δ_D of 0.2, and the analysis scenarios have a demand range up to 0.5. Regarding total costs, RBPM performs still better than the IDPM. However, RBPM does not facilitate all demand, lookiin at the not-facilitated-demand, NFD sometimes reaches 10% of the total demand.

Figure 6.12 zooms in on the behaviour of RBPM per different demand range. As one can see, the demand ranges of 0.3, 0.4, and 0.5 are the main cause of the not facilitated demand outliers. Which is as expected as these demand range scenarios are not used as optimization scenarios. The kde plots are not a one-on-one representation of the counts for the not facilitated demand as the lines are smoothed out in order to interpret the results better. The exact counts of the demand not facilitated per demand range can be seen in Figure I.10. That Figure portrays the same behaviour, and the values do not go below zero. Something that appears to be the case in the kernel density plots.

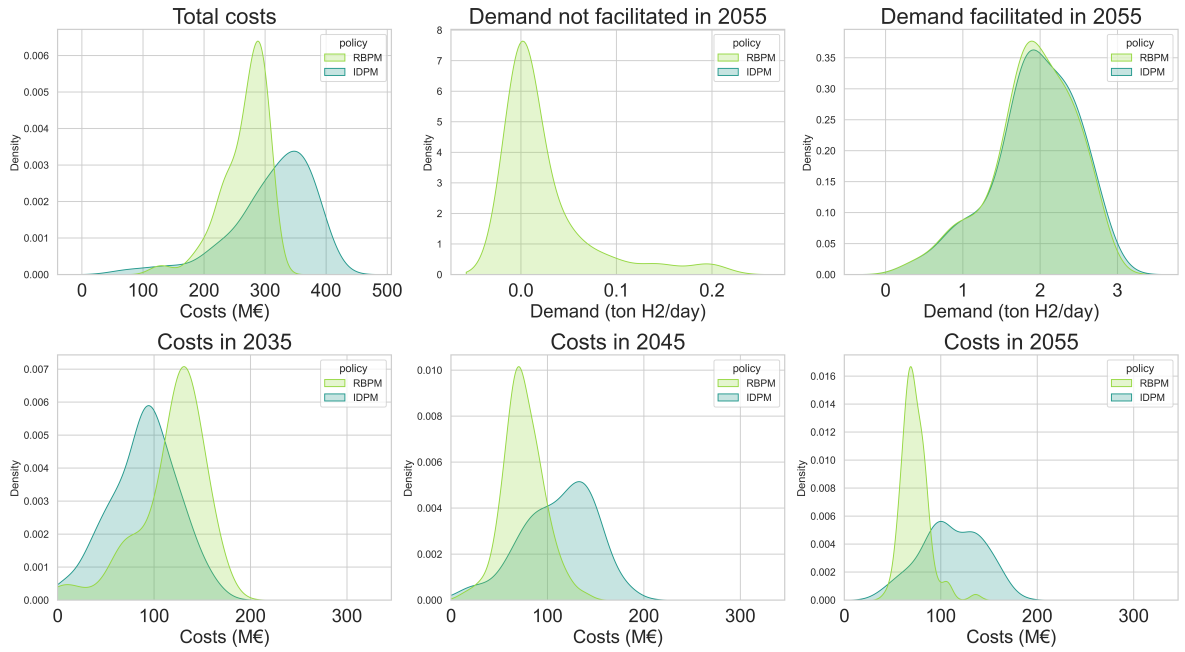


Figure 6.11: Kernel density estimation comparing RBPM and IDPM. The experiment is conducted with higher uncertainty in the analysis scenarios compared to the optimization scenarios. Notably, total costs of RBPM are less compared to IDPM, however RBPM does not facilitate all demand.

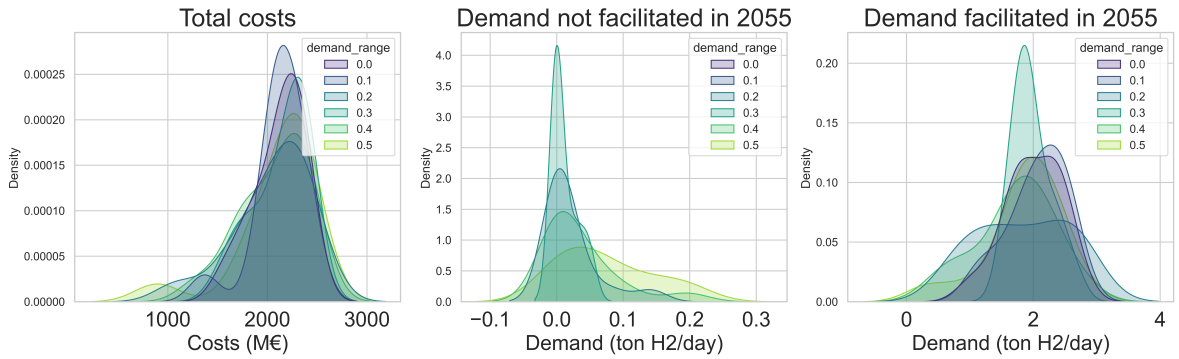


Figure 6.12: Kernel density estimations comparing the behaviour of RBPM for different demand ranges. The experiment is conducted with higher uncertainty in the analysis scenarios ($\Delta_D = 0.5$) compared to the optimization scenarios ($\Delta_D = 0.2$). Notably for higher demand ranges the RBPM does not facilitate all demand.

6.4.2. Optimizing with high uncertainty

Figure 6.13 shows the kernel density estimation plots of the robust network optimized for demand scenarios with demand range 0.5. The performance of the robust network in terms of demand not facilitated and demand facilitated per demand range of the analysis scenarios is much better compared to the previous experiment. All demand except for one scenario is facilitated independent of the demand range, see Figure 6.13.

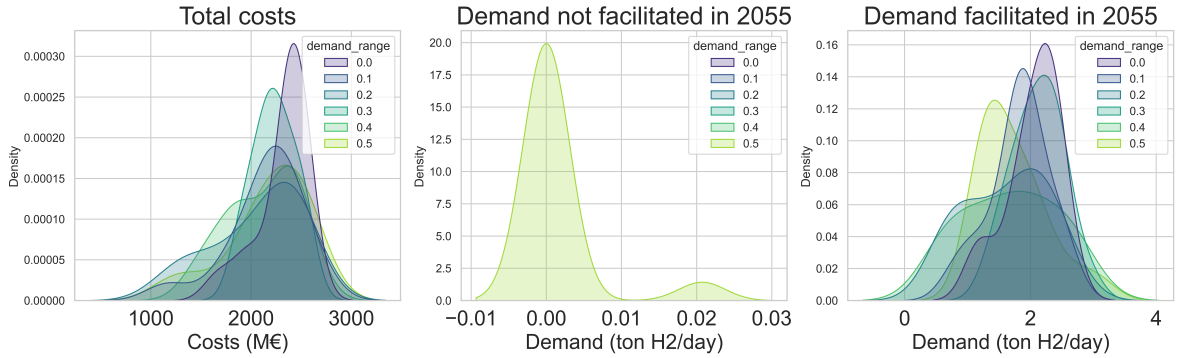


Figure 6.13: Kernel density estimations comparing the behaviour of RBPM for different demand ranges. The experiment is conducted with high uncertainty in optimization scenarios ($\Delta_D = 0.5$). Notably, the demand is facilitated for all analysis scenarios except one scenario with ($\Delta_D = 0.5$

Comparing the high uncertainty optimized RBPM with the IDPM in Figure 6.14, the network on average still performs better in total costs, although, compared to the robust network optimized with $\Delta_D = 0.2$, the total costs are higher for the network optimized with $\Delta_D = 0.5$.

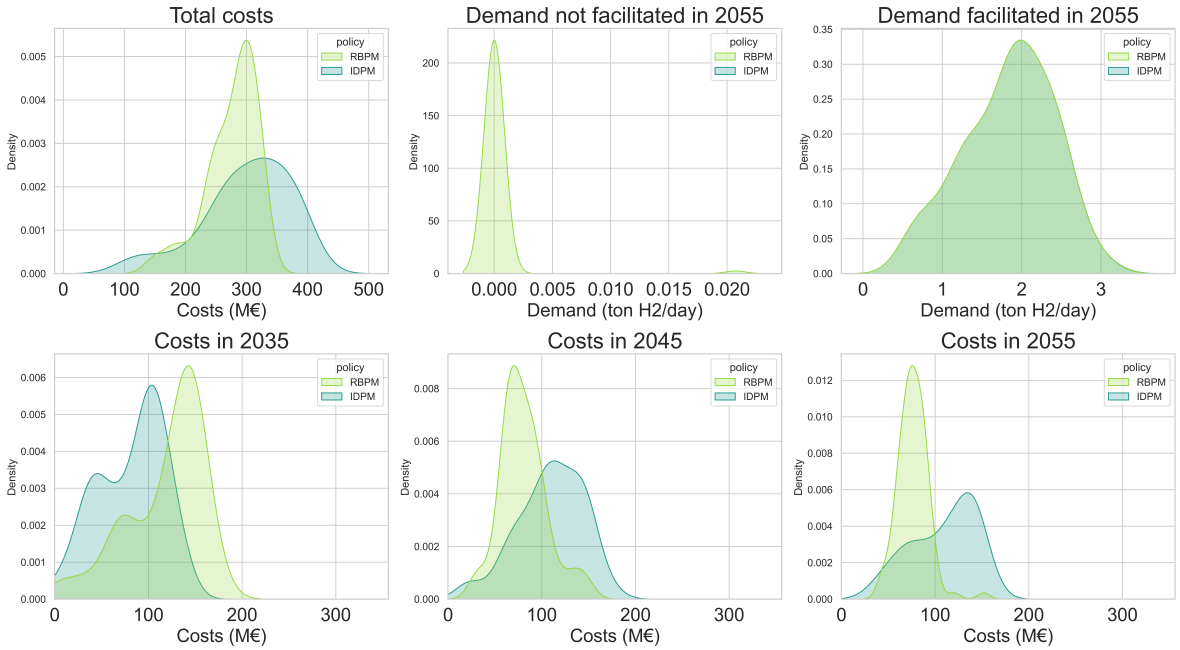


Figure 6.14: Kernel density estimation comparing RBPM and IDPM. The experiment is conducted with high uncertainty in optimization scenarios ($\Delta_D = 0.5$). Notably, total costs of RBPM are less compared to IDPM.

7
Discussion

In this chapter first, the results are interpreted per experiment, and then the limitations of this study are discussed.

7.1. Results interpretation

This section presents the key findings from the study, first by studying the proof of concept and then interpreting experiments 1, 2, and 3.

7.1.1. Proof of concept

The proof of concept showed the robust backtracking plannings method (RBPM) implementation for one configuration, simulating a real life application of the RBPM. The proof of concept also showed how the robust network (RN) topology is derived from the occurrence of edges. And hereby it highlighted, the effect the optimization scenarios have on the RN topology. The RN topology is derived from the in 2055 over time generated networks with IDPM. These IDPM networks in 2055, are a result of the path dependent network optimization per decade and are therefore not the cost optimal network for the demand in 2055. This has advantages and disadvantages.

The advantage is that the RN optimization indirectly distinguishes between nodes that are likely to transition early and nodes that are likely to transition late. While the RN in total might be costlier if all demand nodes are participating, if the demand nodes participate that are most likely to, the RN implementation is probably more cost efficient. For example, the extra nodes in a RN are often connected to the network with long trajectories connected them to the 'backbone' of the network. These trajectories are more costly when they are needed to be placed, however, the benefit is that if they are not placed it is economically more profitable than redirecting the whole network.

However, two high-transitioning nodes transition more frequently across different decades since each transition probability is lower than 45%. The cycles that appear in the occurrence graph, as seen in Figure 6.2, are caused by the differing participation timings of at least two demand nodes. As these nodes transition asynchronously more often, the edges chosen represent the two separate optimized capacitated paths between the demand nodes and the supply node, rather than the combined optimized path or the two transitioning nodes.

This disadvantage is caused by how cycles are dealt with in the RN topology heuristic where the lowest occurring edge is removed to remove the cycle. These cycles can occur frequently in the IPCs due to the grid structure of the road network; where multiple almost as short routes exist. Further research could explore various topology heuristics or conduct a comparative analysis of the derivation of the RN from the scenario-optimized network over time, in relation to the directly cost-optimized networks for the demand of 2055.

Lastly, when comparing the costs to existing hydrogen projects, the estimated costs presented in this thesis are within the same order of magnitude; however, they are slightly higher. For instance, the *HyTransPortRTM* project, which is developing a hydrogen pipeline in the Port of Rotterdam, spans 30 kilometers and has a capacity of 3 ton H_2/ day (CES Rotterdam-Moerdijk, 2021; IEA, 2024b). This project has estimated investment costs of approximately 100 million euros, with a possible variation of

$\pm$ 40%. The estimated costs for the scenario outlined in the proof of concept for the year 2035 range between 120 and 160 million euros.

Recognizing that the proof-of-concept network is around halve the length than that of the *HyTrans-PortRTM* project, the cost estimates in this thesis align with the magnitude of those from existing hydrogen projects, but are higher than what has been announced by these projects. This discrepancy can partly be attributed to the inclusion of operational costs in this thesis, which account for 20% of the total costs projected for 2035. Additionally, it is possible that simplifications in the calculations of diameter to capacity ratios have led to an overestimation of costs, thereby resulting in the higher values reported here.

7.1.2. Comparison with simpler methods

For this experiment, the RBPM is compared with ML-RBPM and IDPM. They are compared by looking at both Iron and Steel and Pulp and Paper clusters for three different layouts. In this section first the results of the ML-RBPM comparison are discussed, and then the IDPM comparison.

Comparison with ML-RBPM

The difference between ML-RBPM and RBPM is how the robust network topology is derived. Namely, for ML-RBPM, the minimum length robust network (ML-RN) topology heuristic is used and for RBPM the robust network (RN) topology heuristic is used.

Looking at costs the RBPM performs compared to ML-RBPM better for every configuration. This outcome aligns with theoretical expectations, as the RBPM is designed to minimize costs, whereas the ML-RBPM minimizes length. There are notable differences in total costs for the ML-RBPM across different configurations. This is attributed to the diverse placement of supply nodes, which significantly affects the RBPPM topology of the network. In contrast to the RBPM, the performance of the ML-RBPM is not influenced by the position of the supply nodes; rather, it is determined exclusively by the IPC layout and the placement of the terminal node. Thus, when the supply node is placed such that the RBPM topology looks more like the ML-RBPM, the methods perform more the same.

An interesting thing is that the full ML-RN topology was sometimes longer in length than the RN topology, see appendix I.1.2. This is due to the fact that the ML-RN topology derivation also uses a heuristic, namely the distance network heuristic (DNH) to derive the MStT. The DNH approximates the MStT by first translating the steiner tree is translated into a graph with only the terminal nodes, where the edges between the nodes represent the shortest paths between the terminal nodes in the steiner graph. Then the minimum spanning tree is selected from this graph and the edges are translated back to the steiner graph.

The DNH has a 'global' approach of generating the MStT; where most optimization is done in the non-steiner graph domain. The edge turn tree heuristic used in the developed method to approximate the minimum capacitated steiner tree (MCStT) , starts with a MStT and iterates over different paths to look for lower costs considering all possible node changes including steiner nodes. Thereby the edge turn tree heuristic has a more local approach. Since the costs are also dependent on length, it is possible that by iterating over different networks and selecting the least costs network, also a network with minimum length is found. If for the ML-RN another heuristic was used that also iterates more, it is probable that a better minimum length was found for some ML-RNs. However, since no heuristics can guarantee the optimal solution this cannot be guaranteed.

However, even if the ML-RN topology is smaller in total length as is the case for Iron and Steel configuration 3, when implemented the RBPM was on average smaller in length than ML-RBPM. This is also inherently caused by the ML-RN topology heuristic and its lack of considering plant specifics. Because, while it can be the case that the ML-RN has minimal length if all demand nodes participate, the ML-RN topology has not considered what happens if some nodes do not participate. While the RN has this aspect way less as more of a backbone with high capacity is created where different nodes are connected to with branches, this increases the total length, but the moment one of these nodes does not participate the length is decreased.

The study conducted by André et al. (2013) involved a comparison between a minimum capacitated spanning tree heuristic and a minimum spanning tree heuristic. The report indicated that the Minimum Spanning Tree (MST) method was 4% shorter in length than the Minimum Cost Spanning Tree (MCST),

but it also resulted in an additional cost of 18%. The analysis of costs aligns with the findings from this experiment. However, the observations related to length are inconsistent with the observations of this thesis.This difference in observation could be connected to the distinction between a minimum steiner tree in a graph problem (MStG) and a minimum spanning tree problem (MST). While MST algorithms always find the optimal solution, MStG algorithms do not.

Comparison with IDPM

This section will compare the RBPM with the IDPM. The relative costs plot in Figure 6.8 demonstrated that, over a 30-year period, the total costs for the RBPM are lower than those for the IDPM across almost all clusters and configurations. However, it also highlighted that the RBPM requires significantly higher investment costs in the first decade, as substantial capacity needs to be installed upfront.

This high investment to implement RBPM in 2035 could be discouraging for decision-makers, who may hesitate to invest heavily in hydrogen infrastructure during the early stages when the future of the hydrogen economy is most uncertain. This is particularly true for clusters with high temperature demands, as the absolute difference between the IDPM and the RBPM is higher, and thereby making it more difficult to finance the budget.

Additionally, there is a difference in performance between low and high demand clusters. Looking at the difference in total costs in 2055, the RBPM performs better in a low total potential demand cluster (Pulp and Paper) than in a high total potential cluster (Iron and Steel); more or equal money is saved when RBPM is implemented for Pulp and Paper compared to Iron and Steel. While in absolute values the cumulative costs for Iron and Steel are higher. Thus, looking at percentages relative more costs can be saved if the RBPM is implemented for networks that in total have less costs.

The lower relative costs for Iron and Steel cluster is caused by the fraction f_{OC} determining operational costs and plant specifics of the Iron and Steel configuration. It looks like there is a tipping point, where at one point the total costs in 2035 are so high that RBPM is eventually as expensive as the IDPM, especially if not all nodes transition towards 2055. This tipping point is almost reached in layout 2 of Iron and Steel where the distances and potential demand are the highest. Because the high installation costs also result in high operational costs per decade, as these are calculated as 40% of the total investment costs per decade, which influence the costs in 2045 and 2055 as well.

Furthermore, the dominant processes of a cluster also influence the cost difference between IDPM and RBPM. For the RBPM capacity is installed for almost all participating nodes, as the RBPM has robust capacity. If one node with a significant potential hydrogen demand does not participate in an analysis scenario, this capacity is already installed by the RBPM if these pipelines need to be used to facilitate other plants. However this capacity will in that scenario not be installed by IDPM and thereby installation costs for IDPM are spared. This increases the difference in costs between IDPM and RBPM in favor of IDPM.

7.1.3. Comparing RBPM performance across different clusters

This experiment looked at the different patterns in costs for the RBPM implementation in different clusters. Looking at the costs per decade for each cluster, two patterns can be visible separating the clusters in two groups.

The first group consist of Iron and Steel, Refineries, and Basic chemicals. This group has higher costs per decade compared to the other group and specifically the costs for 2035, $C(2035)$, are high. The clusters in this group correspond to the clusters with hydrogen feedstock decarbonization processes, and these clusters have the three highest potential demand and have a high transition probability. However, the reason why the $C(2035)$ is so high for this group is not because of the high transition probability, since the Glass cluster (also predominantly high transition probability) does not have the same cost pattern as this group. Therefore, the reason of the high costs, especially the $C(2035)$, is the high potential demand.

The second group consists of the other five clusters; Metal processing, Non-ferrous metals, Glass, Mineral processing, and Pulp and paper. This group still has the same pattern in costs per decade, where $C(2035)$ is the highest compared to $C(2045)$ and $C(2055)$. Only in the proportion of $C(2035)$ compared to 2045 and 2055 decades does this group differ from the first group, as $C(2035)$ is not twice

as much as $C(2045)$. $C(2035)$ being not much more than $C(2045)$ is because the robust capacity that needs to be installed is less for the second group clusters.

For decision makers, it is a good sign that the clusters with low hydrogen potential, who are most economical preferable for RBPM implementation, also have on average low costs in 2035, as this makes it easier to invest in a robust network, as less absolute money needs to be invested, compared to the feedstock group.

The costs per decade for the RBPM show that configurations significantly affect the cost per decade pattern, especially regarding the projected costs of $C(2045)$ compared to $C(2035)$. This is primarily observed in clusters where decarbonization technologies have a low technology readiness level (TRL) or low probability distribution. In these cluster, the randomly selected other 50% of the plants have a greater influence on the hydrogen demand in 2035 and, consequently, the associated costs of the network. For example, if these other 50% of the plants consist largely of high-demand facilities, such as iron and steel, basic chemicals, or refinery plants, this results in an increase in $C(2035)$. Conversely, if these plants have low potential demand or a TRL below 6, $C(2035)$ will decrease.

7.1.4. Performance under demand uncertainty

The RBPM showed sub optimal performance in scenarios characterized by higher uncertainty than trained for. This is caused by the under representation of the estimated demand generated within the high demand range in the optimization scenarios. This is demonstrated by the over representation of analysis scenarios with demand ranges of 0.4 and 0.5 among the outliers of unmet demand. This observation shows that a robust network is only as robust for the scenarios that are generated, a finding corroborated by Beh et al. (2017) in their development of a cost-efficient water infrastructure.

Training the RBPM on optimization scenarios with greater demand uncertainty, characterized by a wider demand range, improved the RBPM performance on facilitated demand. However, this came at the cost of reduced cost efficiency. This aligns with Bartholomew and Kwakkel (2020) observation that when a solution becomes more robust across multiple scenarios, it becomes less fit regarding costs for individual scenarios.

This shows the trade-off this thesis started with; to facilitate all demand a network is very costly. However, from this experiment, it is concluded that it is wiser to consider a littler wider uncertainty range than a smaller uncertainty range. However, this broader uncertainty should be considered within limits, as excessive uncertainty may render the RBPM overly robust across multiple scenarios, leading to suboptimal cost performance in many individual scenarios when compared to the IDPM.

This experiment also highlights an important difference between the IDPM and RBPM implementation. IDPM imposes the condition that demand must be satisfied at every timestep in every scenario, whereas RBPM does not have this requirement. Rather, RBPM is the result of robust optimization over the optimization scenarios using the 90th percentile robustness metric. Thereafter the RBPM is implemented without adhering to the constraint of demand facilitation; when capacity does not satisfy demand, no pipeline is expanded for RBPM.. This outcome highlights the challenges of directly comparing IDPM with RBPM; should RBPM adhere to the same constraints as IDPM, its costs would rise, however, it is unlikely that these costs would reach the level of IDPM.

7.2. Limitations

In this section, the limitations of this study are discussed. The limitations are addressed in the following order: first, the limitations of the IPC model; second, the limitations of the demand scenarios; third, the discussions on deep uncertainty; and lastly, the limitations in the network optimization.

7.2.1. IPC model

The IPC model, while effective for illustration and method exploration, lacks relevance to the real life industrial ports as there are a lot of assumptions regarding the IPC model. Below the assumptions regarding wet infrastructure, industrial node placement, and the generation of IPC configurations are discussed.

Wet infrastructure

First of all, the wet infrastructure does not fully resemble real wet infrastructure. Although, the water infrastructure is mostly selected by having a visual similarity with European ports, it is not possible in the model to let the river flow through the full IPC area, as is the case for example in the Port of Antwerp. Another limitation in the wet infrastructure is that the real life docks differ in width and its approaching channels and docks are more straight that the wet infrastructure used in this thesis. As a result the water ways are more random in the developed layout than in real life is the case.

Two other assumptions made regarding the IPC layout is that the road network cannot cross the water, and that the pipelines will always follow the road network. These two assumptions limit the possible hydrogen networks greatly and make the hydrogen network composition completely dependent on the water infrastructure layout. Furthermore in real life, these two assumptions do not hold since for the ports that are crossed by a river, these assumptions would make it impossible for two halves of the port to share infrastructure. Therefore, in real life it is unlikely that this is the case, and it is likely that there are a few places where roads and pipelines will cross the water, by tunnel or bridge.

Industrial node placement

Another important aspect of the IPC layout is the placement of the industrial nodes in relation to the docks and the IPC size. It is assumed that the industrial nodes are very close to the actual docks. Specifically, for the small layouts 0 and 1, the industrial nodes are packed together all within a radius of one kilometer of the water.

In reality, it might be more likely that the industrial plants are located near the port, but within a range of multiple kilometers of the water. This would imply that a port would look more like layout 2; where some industrial plants are located further away from the water, and as a consequence the road network is more extensive and therefore, there are more possibilities for the hydrogen network to form.

However, one must keep in mind that when bigger IPC areas are considered with less area occupied by water, the assumption of pipelines following the road network should be revisited. As less water occupation results in less constraints for the pipelines, and therefore it is more likely that pipelines will cross land where there are not roads.

Reassignment natural gas pipelines

Natural gas pipelines are already placed in industrial areas, to provide heat and feedstock to the industrial plants. It is expected that an extensive amount of natural gas pipelines will be unnecessary towrds 2050 as fossil fuel used will need to be limited and there are multiple studies and case studies in the reassignment of natural gas pipelines to hydrogen pipelines (IEA, 2023a). The reassignment of natural gas pipelines is not considered in this thesis, but can be a fruitful addition for the method. However, it is important to note that the reassignment of natural gas pipelines regarding safety, operation, and conversion costs is still an on-going academic debate (Martin et al., 2024). Therefore, the costs for reassignment and the operation costs of these old natural gas pipelines should first be studied before this can be implemented. However, once the reassignment cost function is known the implementation into the existing robust method is straight forward as most of the code is already made by Heijnen et al. (2019).

IPC configuration

An IPC configuration is classified as a cluster of a specific industrial subsector if at least 50% of the plants within that configuration belong to that subsector and the rests of the plants subsectors are arbitrary assigned. However, the threshold of 50% for the cluster coefficient is an assumption that lacks empirical or academic validation. In contrast, European Observatory for Clusters and Industrial Change (2020) identified 24 clusters in the province of East Flanders, which includes part of the Port of Antwerp. Although not all of these are classified as industrial clusters, this observation suggests that various subsector clusters can coexist within the same geographical region. While this thesis acknowledges some heterogeneity among the plant subsectors, it does not assume that multiple subsectors can coexist within the same IPC. As a result, more homogeneity or heterogeneity might be assumed within the IPC configuration than might be realistic.

Despite these limitations, the IPC model serves as a valuable tool for testing the robustness of the RBPM method. It enables a preliminary exploration of how the methodology can be applied across

various network configurations. Moving forward, it would be interesting to investigate the method's performance in real-world industrial port clusters to validate the method. Furthermore, it would be interesting to look for other aspects defining these industrial port clusters, other than the cluster coefficient.

7.2.2. Demand scenarios

There are numerous simplifications in the generation of demand scenarios. The first simplification is that the transition uncertainty parameter is based on the presence of (prospected) competing technologies. However, it does not take into account interdependencies between plants. Specifically, in each time step, a plant independently 'decides' whether to transition, without considering the actions of other plants or the availability of infrastructure. This fails to reflect reality, where industrial clusters typically exhibit a high degree of interdependence, with some plants relying on each other's by-products which are influenced by process change of neighbors (Cioli et al., 2021). Consequently, when one plant transitions, it can have a cascading effect on others, which is not modeled in this thesis.

Secondly, the estimated demand is derived from projected demand, calculated based on the production capacity of the plants and a demand range that allows for deviations. This assumes a binary demand behavior, whereas, in reality, it could follow a stepwise pattern where part of the production transitions first, followed by additional increments later. Furthermore, for simplicity, it is also assumed that one node represents only one process; in reality, a plant can execute multiple processes while having only one connection to the network. As a result, the demand pattern per node modeled is simpler than it is in reality. A stepwise demand function considering multiple processes per plant would expand the range of plausible future scenarios, making the robust optimization problem more complex, as each plant would have more potential demand outcomes by 2055.

7.2.3. Deep uncertainty

These simplifications in demand scenarios raise questions to which extend the robust approach addresses deep uncertainty of the system. For instance, what would occur if hydrogen for high-temperature processes fails to kick-off? This thesis does not consider that potential future. There exists a trade-off between constraining the future solution space to reduce computation time and encompassing all potential futures to develop a highly robust network. In this thesis, the trade-off is established as follows: participant uncertainty is defined within a range for each plant, while demand uncertainty is similarly constrained within a range of an estimated demand, resulting in limitations on potential future demand scenarios. Thereby there is a subset of the plausible futures generated, which places the demand scenarios in the middle of exploratory and predictive modeling scenarios (Maier et al., 2016). Therefore, the presented results do not capture the deep uncertainty that the system is. Thereby the generated robust network is only robust for a certain range of plausible futures.

While the choice of parameter space has resulted in findings that do not account for deep uncertainty, the methodology itself has been developed from a deep uncertainty framework. However, it remains untested whether the RBPM will yield the same results under deep uncertainty. When deep uncertainty is considered, uncertainty parameters are more broadly defined; for instance, the high transition probability range is set to [0, 50] instead of the narrower range of [0.3, 0.45]. Consequently, scenarios in which high-temperature processes fail to kick-off is plausible within these deep uncertain scenarios.

Moreover, one potential outcome of incorporating deep uncertainty is that the total costs associated with the robust network may exceed those of the IDPM. This phenomenon arises from the selected robustness metric, which dictates that the 90th percentile of demand must be facilitated over all optimization scenarios. As a result, the robust capacity installed in the network will be high, even when the transition probability is low. Consequently, in certain low-transition scenarios, the IDPM may outperform the robust network. This behavior could not be observed in the current study due to the constraints imposed on the parameter space, which were implemented to limit computational time.

7.2.4. Network optimization

In this section the limitations of the cost function will be discussed. Furthermore, the network optimization method minimizes the costs by altering the pipeline locations and capacity. However, this method does not consider specific diameter sizes of pipelines nor does it takes into account fluid flow or uncertainty in the costs over time.

Cost function

The cost function used in this method is adapted from Hammond et al. (2024). They utilized the cost function in a local industrial pipeline network, however its application was slightly different as it was used in a MILP. In this study, the cost function is defined on a per-edge basis instead of a per-straight pipeline basis, which creates a problem for the cost function. For instance, if a pipeline extends directly for 600 meters but intersects three nodes in the model, it is regarded as placing a pipeline three times, as the pipeline consists of three edges. This presents an issue as the cost calculation (Eq 4.8) includes a constant that does not depend on length nor diameter. This constant is approximately 0.5 million euros and is connected to one time costs independent of length or capacity (Hammond et al., 2024). This limitation results in increased costs for each network and favors long edges without interruption of steiner nodes, rather than pipelines that run straight. Nevertheless, since the cost function is consistent across all implementations and networks, it is unlikely to significantly influence the comparison when the same layout is used.

Furthermore, construction costs may increase over the decades, a factor not considered in the current cost function. While this thesis does not address changes in construction costs over time, it is important to note that such changes will influence overall costs, particularly the costs per decade for IDPM, since the implementation of IDPM requires the continual expansion of network trajectories to enhance capacity every decade. It is anticipated that pipeline construction projects will also decarbonize by 2055; however, the timeline for the decarbonization of construction projects remains uncertain due to various external drivers (Arogundade et al., 2023). This uncertainty introduces a variable into the cost function that has not been addressed, as it was outside the scope of this study. Furthermore, if construction equipment is not decarbonized, the potential for greenhouse gas emissions could result in future penalties, affecting financial considerations. Conversely, it is possible that installation costs may decrease as more hydrogen pipelines are established, benefiting from economies of scale. Additionally, inflation is not accounted for in the cost function; if it were, it is conceivable that the performance of IDPM would be lower compared to RBPM.

Pipeline diameter

The developed method determines the capacity of a pipeline without any constraints on its diameter. However, in reality, only a limited range of diameters may be available (Johnson & Ogden, 2012; Reuß et al., 2019). The restriction to a few diameter options can significantly affect both the IDPM and RBPM. Implementing the diameter constraints may unintentionally make the planning methods either less robust or more expensive, as restricting diameter options results in incremental behaviour in both cost and facilitated demand per edge, rather than achieving a smooth optimization. This may dictate the network to implement sub-optimal yet viable solutions that satisfy demand at increased costs and with surplus capacity.

Fluid flow dynamics

Secondly, in this method, the detailed flow characteristics of pressurized hydrogen is not considered as only a rough estimation is made when linking pipeline diameter to capacity. This is a significant limitation, as the actual flow dynamics—especially in high-pressure environments—can lead to issues such as vibrations, which affect the structural integrity of the pipeline over time and thereby influence the safety of pipelines (Raj et al., 2024). Moreover, the pressure drop caused by the fluid flowing through the pipeline is also not taken into account, as a result, pumping stations are not adequately taken into account for the cost function.

However, Raj et al. (2024) also emphasize that a focus on demand scenarios is more important when considering the network, than considering the fluid flow, as incorporating fluid flow only decrease the costs with 2%. Furthermore, since it is important that this thesis is computational fast to compute multiple scenarios, adding full fluid flow dynamics would counter act the possible fast usage of this method. However, it can be beneficial for further research to look at some easy to implement objectives based on fluid flow dynamics, such as limiting the amount of curves a pipeline network takes, since pressure drop is minimal for a straight horizontal pipe (Nandagopal, 2022).

Further research could address the cost function and curves issue by allowing the method to recognize specific edges as components of a continuous pipeline, thereby increasing the applicability of the cost function and additionally align with one of the proposed fluid flow objectives to promote straight pipelines.

8

Conclusion

This thesis aimed to contribute to the development of hydrogen pipeline networks in an industrial port cluster in order to facilitate the transition towards a near-zero emission society. Existing literature in local hydrogen pipeline networks do not consider deep uncertainty when planning a future hydrogen network, rather networks are developed to fit a few likely scenarios. However due to the complexity of the hydrogen energy system, these few likely scenarios cannot capture the future of such a complex system, as the deep uncertainty of the future makes it difficult to predict the outcome (J. H. Kwakkel et al., 2010). This thesis aims to fill this gap by addressing the question: how to plan over time a cost-effective, robust hydrogen pipeline network for industrial port clusters under uncertainty?

To address this research question, a method is developed that helps plan a network that accounts for participant and demand uncertainty in industrial plants within an industrial port cluster. This method is developed and tested throughout this thesis through six sub-questions. First, each sub-question will be answered individually. Subsequently, the overarching research question will be addressed. Following this, the academic, social, and practical contributions of the thesis will be discussed. Finally, recommendations for future research will be provided.

What characteristics of Industrial Port Clusters should be considered for the hydrogen network development?

This thesis develops a method that can be applied to different IPCs. Therefore, it is important to assess the characteristics of an IPC. An industrial port cluster is defined in this thesis as a population of firms located geographically close together and close to a port, that have economically linked activities and want to cooperate. Firms with economically linked activities, such as a shared common stock of products or re-using residual flows, benefit from a cluster as this reduces transportation costs and stimulates cooperation. Important characteristics of an IPC are cooperation, geographical proximity, and economically linked activities, which are often evidenced by the overrepresentation of one subsector within an IPC.

Multiple industrial subsectors have processes eligible for transitioning to hydrogen to decarbonize. These subsectors include iron and steel, glass, mineral processing, metal processing, basic chemicals, non-ferrous metals, refineries, and pulp and paper. These subsectors either have processes that require high -temperature heat or can transition to processes in which hydrogen can substitute fossil fuel feedstock. However, it remains uncertain when or if these processes will transition to hydrogen.

In this thesis, it is proposed that the likelihood of a process in an IPC transitioning to hydrogen is determined by the competing technologies available, with high-temperature heat technologies having a greater chance of transition. Furthermore, the technology readiness level (TRL) of these technologies will also influence the speed of transition over time. This thesis assigns a transition probability per decade to each process, assuming that by predefining a range for this transition probability, it can capture the range of effects from external variables (economic, political, technological, and social).

Additionally, the physical characteristics of an IPC also influence the hydrogen network, particularly the possible locations for pipeline placement. The presence of water and the high concentration of firms in the IPC result in limited space for infrastructure expansion. Therefore, this thesis assumes that the pipelines will be placed next to the roads, thereby avoiding private property and water.

An actor analysis was conducted, and the main take away is that the hydrogen network should minimize operational and installation costs, while also facilitating current and future users. This approach will benefit both the network operator and hydrogen consumers and producers using the network, as lower network costs enable lower tariff rates.

How can cost-efficiency and robustness be operationalized in a physical network?

Because of this complexity it is important that a network is robust. Robustness is defined as a measure of the adequate facilitation of demand of the network under a range of plausible scenarios (Maier et al., 2016). The demand is uncertain as this depends on multiple factors, thus several plausible future scenarios are generated based on the transition probability of the processes in the IPC to grasp the possible future space.

Moreover, the term adequately in the robustness definition can be presented mathematically by different robustness metrics. These robustness metrics take all the outcomes of the performance metric facilitation of demand per scenario and map them to one value. This robustness metric should be chosen in cooperation with involved actors, since the robustness metric represents the risk-averseness of actors. For this thesis, a quite risk-averse robustness metric is chosen; the 90th percentile minimax regret robustness metric. This choice while not made in cooperation with actors is based on a study on future energy infrastructure in the port of Rotterdam (Cuppen et al., 2021).

The 90th percentile minimax regret calculates per scenario the regret; the difference between the facilitated demand of the scenario and the optimal value. A regret based metric is favorable when combined with a cost objective and risk averse actors. The 90th percentile makes the metric less sensitive for outliers, thereby the network can be made robust, without being dominated by one worst or best case scenario.

The other objective, cost-efficiency, is operationalized by saying that installation and operation costs should be minimized given the constraint that the network should be able to facilitate demand. If a network is very robust and will be able to facilitate the demand for every scenario, this network will be quite costly. This shows a trade-off between costs and robustness.

What is a suitable method to generate a robust, and cost efficient hydrogen network over time?

Current traditional infrastructure planning methods employ a time horizon of 10 years. This traditional approach is defined in this thesis as the Immediate Demand Planning Method (IDPM), where planning extends no further than 10 years into the future. Consequently, with this method multiple pipelines are positioned alongside each other on the same trajectory as demand increases, which can lead to significant additional costs, as installation expenses constitute a big portion of the overall costs.

The method developed in this thesis is the robust backtracking planning method (RBPM), where first a robust network is generated using robust optimization and then over time the robust network is deployed using backtracking. Below, first the robust network generation part shall be described, and then the backtracking method shall be explained further.

Robust optimization

A robust network is derived by first generating multiple demand scenarios using the transition probability of the industrial plants. These demand scenarios represent the demand of the IPC in the years 2035, 2045, and 2055. For each decade the hydrogen network is planned according to the IDPM where the network is planned in such a way that the demand is facilitated for that demand scenario. This network optimization is conducted with an adapted version of the Optimal Network Layout Tool (Heijnen, 2024). Hereby, the calculated pipelines from previous decade is given as input for the next decade. The final scenario optimized networks in 2055 are then used to generate a robust network. In this thesis, two robust network generation are tested, the robust network and the minimum length robust network

The robust network is generated in two steps. First, the robust topology is extracted from selecting the most occurring edges that form a tree. Secondly, the robust topology is given capacity using a heuristic that uses the robustness metric. This heuristic starts with the robust topology with maximum capacity

for each edge, and then subtracts capacity of this edge for as long as the network scores optimal for the robustness metric, the 90th percentile minimax regret.

Backtracking planning method

The robust network is then implemented over time using backtracking. Where each decade only the part of the network are deployed that are needed to facilitate the demand (André et al., 2014). Compared with the IDPM the placed pipelines have more capacity and can thereby also facilitate future demand, as a result the pipelines placed with either the RBPM do not have to be extended every decade.

The RBPM is suitable because hydrogen demand is projected to grow significantly by 2055. By accounting for this anticipated demand growth, estimating it in advance, and then backtracking through incremental planning stages, this method can help reduce installation costs. Given that future demand remains uncertain, a robust approach is essential for approximating demand variations and designing a robust network, which can be implemented using the backtracking approach.

How does the developed method perform compared to simpler planning methods?

The experiments comparing the RBPM performance with the IDPM and minimum length robust backtracking planning method (ML-RBPM) demonstrate that the RBPM is in general economically more beneficial in terms of total costs incurred between 2025 and 2055.

For the ML-RBPM, the minimum length robust network is generated in two nearly identical steps as the RBPM; however, it differs in how the topology is derived. The ML robust topology is generated not by using the scenarios but by calculating the minimum steiner tree, which connects all the industrial nodes. All other steps are the same as those in the RBPM.

The ML-RBPM incurs higher costs every decade relative to the RBPM due to the ML-RN topology heuristic, which does not consider specific configuration details, such as supply node location, or plant transition probabilities. The RBPM indirectly considers these configuration specifics, as it uses scenario-optimized networks in its topology heuristic. Consequently, the RBPM consistently outperforms the ML-RBPM in terms of cost efficiency.

Comparing the IDPM to the RBPM, the IDPM shows lower costs in the first decade. This is due to the high capacity required for the RBPM, which corresponds to the robust capacity calculated for the projected high demand in 2055. The higher initial investment required for the RBPM in 2035 may discourage decision-makers, who might be reluctant to invest heavily in hydrogen infrastructure during the early stages of development, especially given the uncertainty surrounding the future of the hydrogen economy. This is particularly relevant for clusters with high potential demand, where the absolute difference in needed budget is more compared to low demand clusters.

Another potential challenge associated with the implementation of the RBPM in high-demand clusters is that the projected profit from utilizing the RBPM instead of the IDPM decreases as both distance and demand increase. As a result, it may become increasingly difficult for decision-makers to advocate for a larger initial investment and to embrace the associated risks, particularly when there is no guarantee of achieving substantial savings by 2055.

What is the performance of the developed method across different clusters?

The results of the costs per decade when implementing the RBPM across different clusters show two patterns corresponding to two groups of clusters.

The first group includes the Iron and Steel, Basic chemicals, and Refineries clusters. These clusters are characterized by high total potential demand and involve processes that require high-temperature heat or can transition to hydrogen feedstock technologies with a high technology readiness level. The costs projected for 2025-2035 in this high-demand group are, on average, twice as much as the costs for 2035-2045. The reason for this is twofold. First, due to the high transition probability of the plants, many will need to be connected to the grid in the first decade. Second, because the total potential demand is high, the robust capacity that must be installed with the RBPM is significant, thereby increasing the costs of the network.

The second group corresponds with the following clusters; Metal processing, Non-ferrous metals, Glass, Mineral processing, and Pulp and paper. These clusters have lower total potential demand and as a result, the costs in the first decade are lower, but on average still higher than the costs for the second and third decade, however, the difference is not so substantial as for the first group.

Additionally, the second group shows more deviations in RBPM performance per configuration, particularly for clusters characterized by low transition probability processes or by processes where the hydrogen technology readiness level is still in the prototyping phase. As a result, it is highly unlikely that these processes will transition by 2035. The other plants that do not belong to the cluster specialization have a greater effect on the cost behavior of the configuration in 2035. Hereby, the total potential demand of the configuration is a good indication of the costs in 2035 and 2045, where a high potential demand correlates with higher costs in 2035.

How does the developed method perform under different demand uncertainty ranges?

The RBPM showed sub optimal performance in scenarios characterized by higher uncertainty than trained for. This is caused by the under representation of the estimated demand generated within the high demand range in the optimization scenarios. This observation shows that a robust network is only as robust for the scenarios that are generated.

Training the RBPM on optimization scenarios with greater demand uncertainty, characterized by a wider demand range, improved the RBPM performance on facilitated demand. However, this came at the cost of reduced cost efficiency. Since training the performance of RBPM with more demand uncertainty still resulted in a more cost efficient network compared to IDPM, it is advised that in case of doubt more demand uncertainty is used for optimization than less. However, due to computational limits, this is not tested for complete demand uncertainty, so the increase of demand uncertainty should be taken within limits.

8.1. Main research question

How to plan over time a robust and cost efficient network over time in an industrial port cluster under uncertainty?

The best planning method for a network in an IPC where the main goal is to facilitate demand in a cost-efficient manner depends on the IPC characteristics, the available budget, and the demand and cost uncertainties. In this thesis the stakeholders main objective was to facilitate demand in every scenario by developing and maintaining a network with the least amount of costs possible. The future scenarios in this thesis are demand scenarios that have uncertainty in plant participation, and the demand of these plants in an IPC.

A robust backtracking planning method (RBPM) is developed, that first, applies a robust optimization heuristic to synthesize the different scenario optimization networks, and secondly, uses backtracking to deploy the robust optimized network over time for the decades 2035, 2045, and 2055. The RBPM is suitable because hydrogen demand is projected to grow significantly by 2055. By accounting for this anticipated demand growth, estimating it in advance, and then backtracking through incremental planning stages, this method can help reduce installation costs. Given that future demand remains uncertain, a robust approach is essential for approximating demand variations and designing a robust network, which can be implemented using the backtracking approach.

This RBPM is compared to the immediate demand planning method (IDPM). The IDPM allocates capacity for the network every decade, ensuring that the precise amount of capacity is assigned to meet demand while minimizing costs over that decade. In contrast, the RBPM aims to achieve cost efficiency over a three-decade horizon, while meeting demand every decade. This thesis demonstrates that the total potential hydrogen demand and the length of the IPC influence the choice between IDPM and RBPM.

Additionally, the performance of RBPM regarding costs is also affected by demand uncertainty. Notably, when the RBPM was tested with a doubling of demand uncertainty, there was a positive impact on facilitated demand without a substantial increase in overall costs. Consequently, it is recommended to consider a higher level of demand uncertainty for optimization when levels of uncertainty are unclear. However, it is important to note that while increasing demand uncertainty can be beneficial, it should be constrained within reasonable limits due to the computational challenges associated with assessing full demand uncertainty.

The RBPM is the most cost efficient if the IPC has a low total potential hydrogen demand, corresponding with IPCs with a low share of Iron and Steel, or Basic chemical or Refinery plants. The RBPM installs pipelines with robust capacity that are needed to facilitate the demand until 2055, thereby saving the costs placing new pipelines for additional capacity every decade.

The RBPM is less beneficial regarding saved costs until 2055 compared to the IDPM, if the IPC has a high total potential demand, corresponding with an iron and steel, basic chemicals or refineries cluster. This is due to the amount of operational costs. Additionally, the investment costs for RBPM in the first decade are higher, which can cause problems regarding budgeting to install the RBPM network. Therefore, the best planning method of a hydrogen pipeline network in a cluster with a high share of high demand plants, depends on the budget available, the expected rise or decrease of the installation costs and the estimated operational costs.

8.2. Contributions

8.2.1. Academic Contribution

The knowledge gap identified is that, while some studies on hydrogen networks address demand uncertainty and local-scale development over time, no research combines all these elements. In this thesis a robust backtracking planning method is proposed that considers uncertainty in participants and demand to generate a robust network for 2055. Then using backtracking, the planning method resembles the over time deployment of the network per decade. This planning method approach is also implemented by André et al. (2014), where the hydrogen capacitated network over time deployment for North France was calculated for a low and a high scenario. However, André et al. (2014) did not consider uncertainty in demand, nor did they consider constraints or obstacles when planning the net-

work. This thesis has contributed by taking into account uncertainty in demand and participants when calculating the network in 2055 and applying the method to an area with a high amount of constraints and obstacles by approaching the problem as a Steiner tree in a graph problem.

However, this thesis has not filled the literature gap regarding deep uncertainty as this thesis does not consider all potential futures. There exists a trade-off between constraining the future solution space to reduce computation time and encompassing all potential futures to develop a highly robust network. In this thesis, the trade-off is established as follows: participant uncertainty is defined within a range for each plant, while demand uncertainty is similarly constrained within a range of an estimated demand, resulting in limitations on potential future demand scenarios. As a consequence the generated scenarios are a subset of all plausible futures, which places the demand scenarios in the middle of exploratory and predictive scenarios. Where exploratory scenarios are scenarios that cover the full plausible future space and predictive scenarios indicate the most likely future scenarios (Maier et al., 2016). Thereby, the generated method does take into account some likely future uncertainty but not the full possible future space that is required to consider deep uncertainty.

8.2.2. Societal contribution

This thesis contributes to the realization of a near-zero-emission (NZE) society by investigating the development of cost-efficient and robust hydrogen networks in industrial clusters. Hydrogen will play a role in achieving near-zero emissions by 2050. However, as of now, the transmission and demand side of hydrogen development are lagging (IEA, 2023a). By providing insights into how hydrogen networks can be developed efficiently, and showing that early investments in a robust network will result in lower long-term costs, this research shows a beneficial situation in the long term for network developers, and decision makers; Investing now in robust infrastructure will not only reduce overall costs towards 2055, but also address the IEA (2023a) concern regarding the insufficient transmission capacity currently being developed to meet the 2055 NZE targets.

Furthermore, this thesis emphasizes the significance of the temporal deployment of network studies, demonstrating that the optimal planning strategies vary depending on the specific IPC configuration. This study underscores that it is not solely the studies focused on the cost-efficient future end state of the system in 2055 that are important; additionally the over time deployment for achieving that end state over time is equally crucial. Such considerations are essential for creating a future hydrogen network in a cost-efficient manner.

8.2.3. Practical relevance

Although the proposed planning method has not been applied to a real case study, several recommendations can still be made for network developers in industrial port clusters (IPCs). Firstly, while it might seem more attractive to delay investment until market conditions improve and demand for hydrogen increases, it is likely to result in higher long-term costs. Developing hydrogen networks now, particularly within industrial clusters, allows stakeholders to avoid costly capacity upgrades later by building in scalability from the outset. When operational costs are expected to be low or when clusters have moderate or low hydrogen demand, implementing a robust backtracking method can be economically advantageous, providing long-term benefits without significantly higher initial investments.

Secondly, when considering uncertainty ranges for robust optimization, it is advisable to incorporate a higher level of demand uncertainty rather than a lower one in situations where there is doubt. While this thesis did not test the method under conditions of complete demand uncertainty due to computational constraints, it was found that applying increased demand uncertainty within reasonable limit enhances the robustness of network optimization without sacrificing too much in terms of costs.

8.3. Future research

Future research should incorporate the various factors influencing demand dynamics within IPCs to further refine and expand the robust hydrogen network planning method developed in this thesis. This research can include accounting for plant interdependencies, and a stepwise approach to demand growth. Moreover, applying the RBPM to real-world IPCs could offer valuable insights into the factors that characterize an IPC and how these factors influence hydrogen network development, while also helping to validate the method's effectiveness.

Another promising area for future exploration is examining the behavior of the RBPM under conditions of deep uncertainty by relaxing the constraints on uncertainty parameters. This thesis did not address this relaxation due to computational limitations; however, it holds significant potential given the system's deep uncertainty. Future studies could assess the performance of the RBPM in deeply uncertain scenarios and explore the balance between cost-efficiency and robustness. Additionally, these studies could investigate how varying robustness metrics might impact the overall performance of the RBPM.

Exploring alternative approaches to robust optimization also holds potential. For example, future studies could also examine alternative robust optimization approaches, either by tailoring the robustness metric to better align with stakeholder preferences or by testing different heuristics for robust network generation. This would allow for analysis of performance across both high and low demand clusters, potentially leading to a more refined method tailored to high-demand environments.

Moreover, enhancing the RBPM with adaptive strategies would be valuable. Although pipelines are inherently fixed once installed, planning over a 30-year horizon may allow for some level of flexibility. Using cluster analysis of scenario-optimized networks could help identify potential points where adaptive decisions are feasible. This approach would enable decision-makers to defer certain pipeline investments, allowing time to respond to evolving hydrogen demand and make strategic decisions later as circumstances become clearer.

Another important area for investigation is the cost function. Examining how changes in capital and operational expenses over time influence the RBPM and IDPM would offer a more nuanced understanding of cost dynamics. Key considerations include the impact of high initial investment costs, and the proportional influence of operational costs. Additionally, future studies should consider uncertainties in long-term construction costs, such as potential greenhouse gas penalties or cost reductions through economies of scale. A more nuanced understanding of these factors would refine the cost function and enhance the understanding and performance of the RBPM and IDPM over time.

Finally, developing an intuitive decision-making interface would enable a better examination of trade-offs among various objectives and interpretations of robustness, improving both the accessibility and practicality of the RBPM. Future studies could test this interface with actual users to enhance the usability of the RBPM. Additional features, such as natural gas pipeline reassignments and considerations for water crossings, would further increase its relevance in practical, real-world applications.

References

Almansoori, A., & Shah, N. (2011). Design and operation of a stochastic hydrogen supply chain network under demand uncertainty. https://doi.org/10.1016/j.ijhydene.2011.11.091

André, J., Auray, S., Brac, J., De Wolf, D., Maisonnier, G., Ould-Sidi, M. M., & Simonnet, A. (2013). Design and dimensioning of hydrogen transmission pipeline networks. *European Journal of Operational Research*, *229*(1), 239–251. https://doi.org/10.1016/j.ejor.2013.02.036

André, J., Auray, S., De Wolf, D., Memmah, M.-M., & Simonnet, A. (2014). Time development of new hydrogen transmission pipeline networks for France. *International Journal of Hydrogen Energy*, *39*(20), 10323–10337. https://doi.org/10.1016/j.ijhydene.2014.04.190

Arogundade, S., Dulaimi, M., Ajayi, S., Saka, A., & Ilori, O. (2023). Decarbonization of construction projects: A review and interpretive structural modelling of carbon reduction drivers. *Journal of Engineering, Design and Technology*, *ahead-of-print*. https://doi.org/10.1108/JEDT-05-2023-0202

Bartholomew, E., & Kwakkel, J. (2020). On considering robustness in the search phase of Robust Decision Making: A comparison of Many-Objective Robust Decision Making, multi-scenario Many-Objective Robust Decision Making, and Many Objective Robust Optimization. *Environmental Modelling and Software*, *127*. https://doi.org/10.1016/j.envsoft.2020.104699

Beagle, E., Lewis, M., Pecora, B., Rhodes, J., Webber, M., & Hebner, R. (2024). Model to inform the expansion of hydrogen distribution infrastructure. *International Journal of Hydrogen Energy*, *49*, 105–113. https://doi.org/10.1016/J.IJHYDENE.2023.07.017

Beh, E. H. Y., Zheng, F., Dandy, G. C., Maier, H. R., & Kapelan, Z. (2017). Robust optimization of water infrastructure planning under deep uncertainty using metamodels. *ENVIRONMENTAL MODELLING & SOFTWARE*, *93*, 92–105. https://doi.org/10.1016/j.envsoft.2017.03.013

Boja, C. (2011). Clusters Models, Factors and Characteristics. *International Journal of Economic Practices and Theories*, *1*(1), 34–43. https://ssrn.com/abstract=1987198

Bolat, P., & Thiel, C. (2014). Hydrogen supply chain architecture for bottom-up energy systems models. Part 1: Developing pathways. *International Journal of Hydrogen Energy*, *39*(17), 8881–8897. https://doi.org/10.1016/j.ijhydene.2014.03.176

Börjeson, L., Höjer, M., Dreborg, K.-H., Ekvall, T., & Finnveden, G. (2006). Scenario types and techniques: Towards a user's guide. *Futures*, *38*(7), 723–739. https://doi.org/10.1016/j.futures.2005.12.002

Caglayan, D. G., Heinrichs, H. U., Robinius, M., & Stolten, D. (2021). Robust design of a future 100% renewable european energy supply system with hydrogen infrastructure. *International Journal of Hydrogen Energy*, *46*(57), 29376–29390. https://doi.org/10.1016/j.ijhydene.2020.12.197

Cayley, A. (1857). On the theory of the analytical forms called trees. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *13*(85), 172–176. https://doi.org/10.1080/14786445708642275

CES Rotterdam-Moerdijk. (2021). CES Rotterdam-Moerdijk (2021). Retrieved March 21, 2024, from https://www.portofrotterdam.com/sites/default/files/2021-10/CES-rotterdam-moerdijk.pdf

Cioli, M., Schure, K. M., & van Dam, D. (2021, March). Decarbonisation options for the Dutch industrial gases production. https://www.pbl.nl/en/publications/decarbonisation-options-for-the-dutch-industrial-gases-production

Cuppen, E., Nikolic, I., Kwakkel, J., & Quist, J. (2021). Participatory multi-modelling as the creation of a boundary object ecology: The case of future energy infrastructures in the Rotterdam Port Industrial Cluster. *Sustainability Science*, *16*(3), 901–918. https://doi.org/10.1007/s11625-020-00873-z

Dam, K. H. v., Nikolic, I., & Lukszo, Z. (2013). *Agent-Based Modelling of Socio-Technical Systems*. Springer Netherlands. https://doi.org/10.1007/978-94-007-4933-7

Deloitte. (2023, April). Assessment of green hydrogen for industrial heat. Retrieved April 25, 2024, from https://www2.deloitte.com/content/dam/Deloitte/us/Documents/Advisory/us-advisory-assessment-of-green-hydrogen-for-industrial-heat.pdf

Deltalinqs. (2024). About Deltalinqs. Retrieved August 12, 2024, from https://www.deltalinqs.nl/paginas/openbaar/over-deltalinqs/about-deltalinqs

Edmonds, B. (2017). Different Modelling Purposes. In B. Edmonds & R. Meyer (Eds.), *Simulating Social Complexity, Understanding Complex Systems*. Springer International Publishing. https://doi.org/10.1007/978-3-319-66948-9_4

Efthymiadou, M. E., Charitopoulos, V. M., & Papageorgiou, L. G. (2024). Optimal hydrogen infrastructure planning for heat decarbonisation. *Chemical Engineering Research and Design*, *204*, 121–136. https://doi.org/10.1016/j.cherd.2024.02.028

Erdoğan, A., & Güler, M. G. (2023). Optimization and analysis of a hydrogen supply chain in terms of cost, CO2 emissions, and risk: The case of Turkey. *International Journal of Hydrogen Energy*, *48*(60), 22752–22765. https://doi.org/10.1016/J.IJHYDENE.2023.04.300

European Council. (2024, April). Fit for 55 - The EU's plan for a green transition. Retrieved April 12, 2024, from https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55/

European Observatory for Clusters and Industrial Change. (2020). European Panorama of Clusters and Industrial Change. Retrieved September 13, 2024, from https://www.clustercollaboration.eu/sites/default/files/news_attachment/european_panorama_2020.pdf

Europese Commissie. (2022, May). REPowerEU. Retrieved April 25, 2024, from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/repowereu-affordable-secure-and-sustainable-energy-europe_nl

Fan, Z., & Friedmann, S. J. (2021). Low-carbon production of iron and steel: Technology options, economic assessment, and policy. *Joule*, *5*(4), 829–862. https://doi.org/10.1016/j.joule.2021.02.018

GasUnie. (2024). Organisation. Retrieved August 12, 2024, from https://www.gasunie.nl/en/organisation

Google. (2024). Google Maps. Retrieved July 5, 2024, from https://www.google.com/maps/

Griffiths, S., Sovacool, B. K., Kim, J., Bazilian, M., & Uratani, J. M. (2021). Industrial decarbonization via hydrogen: A critical and systematic review of developments, socio-technical systems and policy options. *Energy Research & Social Science*, *80*, 102208. https://doi.org/10.1016/j.erss.2021.102208

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX, 11–15. https://doi.org/10.25080/TCWV9851

Hammond, J., Rosenberg, M., & Brown, S. (2024). A Genetic Algorithm-Based Design for Hydrogen Pipeline Infrastructure with Real Geographical Constraints. *Computer Aided Chemical Engineering*, *53*, 631–636. https://doi.org/10.1016/B978-0-443-28824-1.50106-X

Hanto, J., Herpich, P., Löffler, K., Hainsch, K., Moskalenko, N., & Schmidt, S. (2024). Assessing the implications of hydrogen blending on the European energy system towards 2050. *Advances in Applied Energy*, *13*, 100161. https://doi.org/10.1016/j.adapen.2023.100161

Heijnen, P. W., Chappin, E., & Nikolic, I. (2014). Infrastructure network design with a multi-model approach: Comparing geometric graph theory with an agent-based implementation of an ant colony optimization. *JASSS*, *17*(4). https://doi.org/10.18564/jasss.2533

Heijnen, P. W., Ligtvoet, A., Stikkelman, R. M., & Herder, P. M. (2014). Maximising the Worth of Nascent Networks [Number: 1]. *Networks and Spatial Economics*, *14*(1), 27–46. https://doi.org/10.1007/s11067-013-9199-1

Heijnen, P. W. (2024). Optimal Network Layout. Retrieved April 2, 2024, from https://gitlab.tudelft.nl/pheijnen/optimal-network-layout

Heijnen, P. W., Chappin, E. J. L., & Herder, P. M. (2019). A method for designing minimum-cost multi-source multisink network layouts. *Systems Engineering*, *23*(1), 14–35. https://doi.org/10.1002/SYS.21492

Herman, J. D., Reed, P. M., Zeff, H. B., & Characklis, G. W. (2015). How Should Robustness Be Defined for Water Systems Planning under Change? *Journal of Water Resources Planning and Management*, *141*(10), 04015012. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000509

Huisman, R. (2021). Towards a robust european hydrogen network. https://repository.tudelft.nl/record/uuid:4f15082f-41fe-44fd-8263-1c8e3a53ab18

Huntington, D. E., & Lyrintzis, C. S. (1998). Improvements to and limitations of Latin hypercube sampling. *Probabilistic Engineering Mechanics*, *13*(4), 245–253. https://doi.org/10.1016/S0266-8920(97)00013-1

Husarek, D., Schmugge, J., & Niessen, S. (2021). Hydrogen supply chain scenarios for the decarbonisation of a German multi-modal energy system. *International Journal of Hydrogen Energy*, *46*(76), 38008–38025. https://doi.org/10.1016/j.ijhydene.2021.09.041

IEA. (2019). Innovation Gaps – Analysis. Retrieved October 19, 2024, from https://www.iea.org/reports/innovation-gaps

IEA. (2021). Net Zero by 2050 - A Roadmap for the Global Energy Sector. https://www.iea.org/reports/net-zero-by-2050

IEA. (2023a). Global Hydrogen Review 2023. https://www.iea.org/reports/global-hydrogen-review-2023

IEA. (2023b). World Energy Outlook 2023. https://www.iea.org/reports/world-energy-outlook-2023

IEA. (2024a). ETP Clean Energy Technology Guide – Data Tools. Retrieved June 6, 2024, from https://www.iea.org/data-and-statistics/data-tools/etp-clean-energy-technology-guide

IEA. (2024b). Hydrogen Production and Infrastructure Projects Database. Retrieved October 21, 2024, from https://www.iea.org/data-and-statistics/data-product/hydrogen-production-and-infrastructure-projects-database

IPCC. (2022). *Climate Change 2022 - Mitigation of Climate Change* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009157926

Johannesson, P., & Perjons, E. (2021). *An Introduction to Design Science*. Springer International Publishing. https://doi.org/10.1007/978-3-030-78132-3

Johnson, N., & Ogden, J. (2012). A spatially-explicit optimization model for long-term hydrogen pipeline planning. *International Journal of Hydrogen Energy*, *37*(6), 5421–5433. https://doi.org/10.1016/J.IJHYDENE.2011.08.109

Khan, M. A., Young, C., MacKinnon, C., & Layzell, D. (2021). The Techno- Economics of Hydrogen Compression. *Transition Accelerator Technical Briefs*, *1*(1), 1–36. https://transitionaccelerator.ca/wp-content/uploads/2023/04/TA-Technical-Brief-1.1_TEEA-Hydrogen-Compression_PUBLISHED.pdf

Kim, H., Hwang, S.-J., & Yoon, W. (2023). Industry cluster, organizational diversity, and innovation. *International Journal of Innovation Studies*, *7*(3), 187–195. https://doi.org/10.1016/j.ijis.2023.03.002

Koningsveld, M. v., Verheij, H., Taneja, P., & Vriend, H. d. (2023). *Ports and Waterways: Navigating the changing world*. https://doi.org/https://doi.org/10.5074/T.2021.004

Korner. (2015). *Technology Roadmap Hydrogen and Fuel Cells*. Paris. https://www.iea.org/reports/technology-roadmap-hydrogen-and-fuel-cells

Kou, L., Markowsky, G., & Berman, L. (1981). A fast algorithm for Steiner trees. *Acta Informatica*, *15*(2), 141–145. https://doi.org/10.1007/BF00288961

Kruskal, J. B. (1956). On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. *Proceedings of the American Mathematical Society*, *7*(1), 48–50. Retrieved August 26, 2024, from https://www.cmat.edu.uy/~marclan/TAG/Sellanes/Kruskal.pdf

Kwakkel, J. H., Walker, W. E., & Marchau, V. a. W. J. (2010). Adaptive Airport Strategic Planning. *European Journal of Transport and Infrastructure Research*, *10*(3). https://doi.org/10.18757/ejtir.2010.10.3.2891

Kwakkel, J. H., & Pruyt, E. (2013). Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty. *Technological Forecasting and Social Change*, *80*(3), 419–431. https://doi.org/10.1016/j.techfore.2012.10.005

Li, L., Manier, H., & Manier, M. A. (2019). Hydrogen supply chain network design: An optimization-oriented review. *Renewable and Sustainable Energy Reviews*, *103*, 342–360. https://doi.org/10.1016/J.RSER.2018.12.060

Maier, H., Guillaume, J., Van Delden, H., Riddell, G., Haasnoot, M., & Kwakkel, J. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling & Software*, *81*, 154–164. https://doi.org/10.1016/j.envsoft.2016.03.014

Martin, P., Ocko, I. B., Esquivel-Elizondo, S., Kupers, R., Cebon, D., Baxter, T., & Hamburg, S. P. (2024). A review of challenges with using the natural gas system for hydrogen. *Energy Science & Engineering*. https://doi.org/10.1002/ese3.1861

McPhail, C., Maier, H. R., Kwakkel, J. H., Giuliani, M., Castelletti, A., & Westra, S. (2018). Robustness Metrics: How Are They Calculated, When Should They Be Used and Why Do They Give Different Results? *Earth's Future*, *6*(2), 169–191. https://doi.org/10.1002/2017EF000649

Melese, Y. G., Heijnen, P. W., Stikkelman, R. M., & Herder, P. M. (2017). An Approach for Integrating Valuable Flexibility During Conceptual Design of Networks. *Networks and Spatial Economics*, *17*(2), 317–341. https://doi.org/10.1007/s11067-016-9328-8

Microsoft. (2024). Microsoft Copilot. Retrieved October 21, 2024, from https://copilot.microsoft.com/

Moradi, R., & Groth, K. M. (2019). Hydrogen storage and delivery: Review of the state of the art technologies and risk and reliability analysis. *International Journal of Hydrogen Energy*, *44*(23), 12254–12269. https://doi.org/10.1016/j.ijhydene.2019.03.041

Moreno-Benito, M., Agnolucci, P., & Papageorgiou, L. G. (2017). Towards a sustainable hydrogen economy: Optimisation-based framework for hydrogen infrastructure development. *Computers and Chemical Engineering*, *102*, 110–127. https://doi.org/10.1016/j.compchemeng.2016.08.005

Morosini, P. (2004). Industrial Clusters, Knowledge Integration and Performance. *World Development*, *32*(2), 305–326. https://doi.org/10.1016/j.worlddev.2002.12.001

Mortensen, L., Kringelum, L. B., & Gjerding, A. N. (2023). How industrial symbiosis emerges through partnerships: Actors, platforms, and stakeholder processes leading to collaborative business models in port industrial areas. *International Journal of Innovation and Sustainable Development*, *17*, 205–228. https://doi.org/10.1504/IJISD.2023.127927

Najjar, Y. S. H. (2013). Hydrogen safety: The road toward green technology. *International Journal of Hydrogen Energy*, *38*(25), 10716–10728. https://doi.org/10.1016/j.ijhydene.2013.05.126

Namazifard, N., Vingerhoets, P., & Delarue, E. (2024). Long -term cost optimization of a national low-carbon hydrogen infrastructure for industrial decarbonization. *International Journal of Hydrogen Energy*, *64*, 583–598. https://doi.org/10.1016/j.ijhydene.2024.02.324

Nandagopal, N. S., Pe. (2022). Fluid Dynamics. In *Fluid and Thermal Sciences* (pp. 35–100). Springer International Publishing. https://doi.org/10.1007/978-3-030-93940-3_3

Neuwirth, M., Fleiter, T., Manz, P., & Hofmann, R. (2022a). The future potential hydrogen demand in energy-intensive industries - a site-specific approach applied to Germany. *Energy Conversion and Management*, *252*, 115052. https://doi.org/10.1016/j.enconman.2021.115052

Neuwirth, M., Khanra, M., Fleiter, T., Jovicic, M., & Shinde, M. (2022b). Future hydrogen demands from industry transition towards 2030–a site-specific bottom-up assessment for north-western europe. *ECEEE Summer Study Proceedings*, 1463–1474.

Ochoa Robles, J., Giraud Billoud, M., Azzaro-Pantel, C., & Aguilar-Lasserre, A. A. (2019). Optimal Design of a Sustainable Hydrogen Supply Chain Network: Application in an Airport Ecosystem. *ACS Sustainable Chemistry and Engineering*, *7*(21), 17587–17597. https://doi.org/10.1021/ACSSUSCHEMENG.9B02620

OpenAI. (2024). ChatGPT. Retrieved October 21, 2024, from https://chat.openai.com

Organization, W. M. (2024, January). WMO confirms that 2023 smashes global temperature record. Retrieved April 22, 2024, from https://wmo.int/media/news/wmo-confirms-2023-smashes-global-temperature-record

Paredes-Vergara, M., Palma-Behnke, R., & Haas, J. (2024). Characterizing decision making under deep uncertainty for model-based energy transitions. *Renewable and Sustainable Energy Reviews*, *192*, 114233. https://doi.org/10.1016/j.rser.2023.114233

Parolin, F., Colbertaldo, P., & Campanari, S. (2022). Development of a multi-modality hydrogen delivery infrastructure: An optimization model for design and operation. *Energy Conversion and Management*, *266*, 115650. https://doi.org/10.1016/j.enconman.2022.115650

Pivetta, D., Dall'Armi, C., Sandrin, P., Bogar, M., & Taccani, R. (2024). The role of hydrogen as enabler of industrial port area decarbonization. *Renewable and Sustainable Energy Reviews*, *189*, 113912. https://doi.org/10.1016/j.rser.2023.113912

Port of Rotterdam. (2024). Our organisation. Retrieved August 12, 2024, from https://www.portofrotterdam.com/en/about-port-authority/our-organisation

Porter, M. E. (2000). Location, Competition, and Economic Development: Local Clusters in a Global Economy. *Economic Development Quarterly*, *14*(1), 15–34. https://doi.org/10.1177/089124240001400105

Prim, R. C. (1957). Shortest connection networks and some generalizations. *The Bell System Technical Journal*, *36*(6), 1389–1401. https://doi.org/10.1002/j.1538-7305.1957.tb01515.x

QuillBot. (2024). QuillBot. Retrieved October 21, 2024, from https://quillbot.com

Raj, A., Larsson, I. S., Ljung, A.-L., Forslund, T., Andersson, R., Sundström, J., & Lundström, T. (2024). Evaluating hydrogen gas transport in pipelines: Current state of numerical and experimental

methodologies. *International Journal of Hydrogen Energy*, *67*, 136–149. https://doi.org/10.1016/j.ijhydene.2024.04.140

Reuß, M., Welder, L., Thürauf, J., Linßen, J., Grube, T., Schewe, L., Schmidt, M., Stolten, D., & Robinius, M. (2019). Modeling hydrogen networks for future energy systems: A comparison of linear and nonlinear approaches. *International Journal of Hydrogen Energy*, *44*(60), 32136–32150. https://doi.org/10.1016/j.ijhydene.2019.10.080

Rijksoverheid: Emissieregistratie. (2024). Emissieregistratie per bedrijf. Retrieved March 27, 2024, from https://data-preview.emissieregistratie.nl/reports

Rothfarb, B., Frank, H., Rosenbaum, D. M., Steiglitz, K., & Kleitman, D. J. (1970). Optimal Design of Offshore Natural-Gas Pipeline Systems. *Operations Research*, *18*(6), 992–1020. https://doi.org/10.1287/opre.18.6.992

Rouwenhorst, K., Krzywda, P., Benes, N., Mul, G., & Lefferts, L. (2021). Ammonia Production Technologies. In *Techno-Economic Challenges of Green Ammonia as an Energy Vector* (pp. 41–83). Elsevier. https://doi.org/10.1016/B978-0-12-820560-0.00004-7

Savage, L. J. (1951). The Theory of Statistical Decision. *Journal of the American Statistical Association*, *46*(253), 55–67. https://doi.org/10.1080/01621459.1951.10500768

Sivanandam, S., & Deepa, S. (2008). Introduction to Particle Swarm Optimization and Ant Colony Optimization. In *Introduction to Genetic Algorithms* (pp. 403–424). Springer. https://doi.org/10.1007/978-3-540-73190-0_11

Tlili, O., Mansilla, C., Linβen, J., Reuß, M., Grube, T., Robinius, M., André, J., Perez, Y., Le Duigou, A., & Stolten, D. (2020). Geospatial modelling of the hydrogen infrastructure in France in order to identify the most suited supply chains. *International Journal of Hydrogen Energy*, *45*(4), 3053–3072. https://doi.org/10.1016/j.ijhydene.2019.11.006

United Nations. (2015, November). *Paris agreement*. Retrieved March 26, 2024, from https://unfccc.int/sites/default/files/english_paris_agreement.pdf

Unsplash. (2021, November). Photo by Christopher Parker on Unsplash. Retrieved April 11, 2024, from https://unsplash.com/photos/a-train-traveling-over-a-bridge-over-water-W6nA-gj7kp0

v. Mikulicz-Radecki, F., Giehl, J., Grosse, B., Schöngart, S., Rüdt, D., Evers, M., & Müller-Kirchenbauer, J. (2023). Evaluation of hydrogen transportation networks - A case study on the German energy system. *Energy*, *278*, 127891. https://doi.org/10.1016/j.energy.2023.127891

Van den Eynde, S., Audenaert, P., Colle, D., & Pickavet, M. (2022). A construction heuristic for the capacitated Steiner tree problem. *PLoS ONE*, *17*(6). https://doi.org/10.1371/journal.pone.0270147

van Vuuren, D. P., Kok, M. T. J., Girod, B., Lucas, P. L., & de Vries, B. (2012). Scenarios in Global Environmental Assessments: Key characteristics and lessons for future use. *Global Environmental Change*, *22*(4), 884–895. https://doi.org/10.1016/j.gloenvcha.2012.06.001

Vine, D. (2021). Clean Industrial Heat: A Technology Inclusive Framework. https://www.c2es.org/document/clean-industrial-heat-a-technology-inclusive-framework/

Weber, A. C., & Papageorgiou, L. G. (2018). Design of hydrogen transmission pipeline networks with hydraulics [Publisher: Institution of Chemical Engineers]. *Chemical Engineering Research and Design*, *131*, 266–278. https://doi.org/10.1016/j.cherd.2018.01.022

Wei, M., McMillan, C. A., & De La Rue Du Can, S. (2019). Electrification of Industry: Potential, Challenges and Outlook. *Current Sustainable/Renewable Energy Reports*, *6*(4), 140–148. https://doi.org/10.1007/s40518-019-00136-1

Welder, L., Ryberg, D. S., Kotzur, L., Grube, T., Robinius, M., & Stolten, D. (2018). Spatio -temporal optimization of a future energy system for power-to-hydrogen applications in Germany. *Energy*, *158*, 1130–1149. https://doi.org/10.1016/j.energy.2018.05.059

Winter, P. (1987). Steiner problem in networks: A survey. *Networks*, *17*(2), 129–167. https://doi.org/10.1002/net.3230170203

X-Rates. (2024, July). Exchange Rate Average (British Pound, Euro). Retrieved July 31, 2024, from https://www.x-rates.com/average/?from=GBP&to=EUR&amount=1&year=2024

Yáñez, M., Ortiz, A., Brunaud, B., Grossmann, I. E., & Ortiz, I. (2018). Contribution of upcycling surplus hydrogen to design a sustainable supply chain: The case study of Northern Spain. *Applied Energy*, *231*, 777–787. https://doi.org/10.1016/j.apenergy.2018.09.047

Yeates, C., Schmidt-Hattenberger, C., Weinzierl, W., & Bruhn, D. (2021). Heuristic Methods for Minimum-Cost Pipeline Network Design – a Node Valency Transfer Metaheuristic. *Networks and Spatial Economics*, *21*(4), 839–871. https://doi.org/10.1007/s11067-021-09550-9

Žerovnik, J. (2015). Heuristics for NP-hard optimization problems - simpler is better!? *Logistics & Sustainable Transport*, *6*(1), 1–10. https://doi.org/10.1515/jlst-2015-0006

A

AI Statement

During the creation and preparation of this thesis, the author utilized *ChatGPT* and *Microsoft Copilot* for coding assistance, and *QuillBot* for paraphrasing during the writing phase (Microsoft, 2024; OpenAI, 2024; QuillBot, 2024). After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of this thesis.

B

Visual Inspection European Ports

A visual study is conducted to search for physical characteristics of industrial ports. This study is done by visually inspecting the biggest industrial ports according to a report by Deloitte (2023). All pictures presented are screenshots made by Google Maps (Google, 2024). Below the Ports are depicted in a large format. The conclusion of the visual inspection and small figures of the ports of Rotterdam, Duisburg, Antwerp, and Dusseldorf are presented in Chapter 2.

Figure B.1: Port of Rotterdam

Figure B.2: Port of Duisburg

Figure B.3: Port of Antwerp

Figure B.4: Port of Dusseldorf

C

Additional derivation cost function

This appendix includes additional information and arguments for the cost function derivation. The additional derivation of the cost function is presented in Section C.1.

C.1. Diameter as function of capacity

As the attentive reader might have noticed, the costs in equation 4.8 are calculated as a function of the diameter. However, the ONLT method optimizes the capacity of the pipeline (Heijnen et al., 2019). The relation between the capacity of a pipeline depends on various parameters where pressure, temperature and friction parameters are very important as we are talking about the speed at which the hydrogen can be transported from node i to node j.

The capacity q_{ij} in kg H_2/day of one pipeline is given by the fluid flow equation (Khan et al., 2021);

$$q_{ij} = 1.1494 * 10^{-3} \left(\frac{T_b}{P_b}\right) \left[\frac{P_1^2 - P_2^2}{G T_f l_{ij} * 10^{-3} Z f}\right]^{0.5} (d_{ij} * 10^3)^{2.5} * 0.0834 \tag{C.1}$$

where the capacity of the pipeline in kg H_2/day is given as a function of the length l_{ij} of one pipe in m, d_{ij} the inside diameter of the pipe in m, the absolute in- and outlet pressure P_1, P_2, and other constants that are named and defined in the table below.

Table C.1: Pipeline Parameters and Constants

Symbol	Name	Unit	Value	Source
l_{ij}	Pipeline length	m	Variable	
d_{ij}	Pipeline diameter	m	Variable	
P_1	Inlet pressure	kPa	7000	Hammond et al. (2024)
P_2	Outlet pressure	kPa	3500	Hammond et al. (2024)
G	Specific gravity	-	0.0696	Khan et al. (2021)
P_b	Base pressure	kPa	101.352	Khan et al. (2021)
T_b	Base temperature	K	288.700	Khan et al. (2021)
T_f	Average flowing temperature of gas	K	288.15	Khan et al. (2021)
Z	Compressibility factor at average temperature and pressure	-	1.031	Khan et al. (2021)
f	Friction factor	-	0.0094	Khan et al. (2021)

For this analysis, it is assumed that there are no height elevations affecting the pipeline. The inlet and outlet pressures are considered to be 70 bar (7000 kPa) and 35 bar (3500 kPa) respectively, this is consistent with what other sources have used as pressure differences. However, in this model the pipeline lengths are smaller than that of the other papers, therefore it is likely that the real pressure drop is lower. While it is acknowledged that T_f, Z, and f are not constant, and depend on the pipeline material and size, the values for the average flowing temperature of the gas (T_f), the compressibility factor (Z), and the friction factor (f) are adopted from the study by Khan et al. (2021). In that study they approached these factors for a 100km hydrogen pipeline.

Next, equation 4.8 is transcribed into C.4 that is used in the model to calculate the cost function using 4.5.

let

$$a_0 = 0.9590 * 10^{-4} \tag{C.2}$$

$$a_1 = \left[\frac{P_1^2 - P_2^2}{G T_f Z f} \right]^{0.5} \tag{C.3}$$

then the diameter can be defined as a function of the capacity and the length, the two outcomes of the ONLT model.

$$D(l_{ij}, q_{ij}) = \left[\frac{q_{ij} l_{ij}^{0.5} P_b}{a_0 a_1 T_b} \right]^{-2.5} \tag{C.4}$$

In this thesis a standard wall thickness of 0.0 mm is assumed. This is not realistic and when the planning method will be implemented in a real case study, the wall thickness should be altered.

D

Heursitics and Flow diagrams

D.1. Pseudo code: Robust Network heuristics

D.1.1. Robust scenario based topology

```
Start with full steiner network (road network)
Assign occurrence attribute to edges using the S scenario-optimized
    networks
Delete the edges with occurrence of zero
FOR edge in network
  Set edge capacity to maximum capacity of edge over all S

%---Remove cycles----
Indicate cycles in network
WHILE cycles in network
  Remove edge with lowest occurrence in cycle
  Add removed 'edges capacity to other edges in cycle
  Indicate cycles in network

%---Remove loose ends---
FOR node in network
  IF node degree = 1 and node is a steiner node
    Remove edge(node, neighbor)
```

D.1.2. ML-RBPM topology

```
Approximate G a minimum length steiner tree with the NetworX steiner tree
    algorithm* that connects the terminal nodes

Assign capacity to G as if the maximum demand of all terminal nodes should
    be facilitated
```

*uses the distance network heuristic (Kou et al., 1981) (see section Theoretic Framework)

D.1.3. Robust capacitated graph heuristic

```
Start with G a robust topology tree with max capacity, and S demand
    scenarios

Calculate robustness value (RV) G for the S demand scenarios (should be
    optimal)
FOR edge in G:
    WHILE RV stays the same and edge capacity > 0:
        Lower the edge capacity with increment_capacity
        Recalculate RV
```

```
        IF RV changes:
           Add increment_capacity to edge_capacity
```

D.1.4. Backtracking robust network implementation

Start with Robust Network RN derived from robust optimization, and demand scenario for three timesteps for analysis scenario. A implemented network N_1, N_2, N_3 is derived for timestep 1, 2, and 3.

```
FOR each timestep t
    FOR each node n with demand(t):
      Derive path n to sink in RN
      FOR each edge in path n to sink in RN:
        IF edge not in N_t
          Add edge to N_t with capacity of edge in RN
      IF NOT path n to sink in RN:
        Generate MCStT with RN and demand(t) as input
        Derive path n to sink in MCStT
        FOR each edge in path n to sink in MCStT:
          Add or update capacity of RN with capacity MCStT
```

D.2. Flow diagram demand scenario

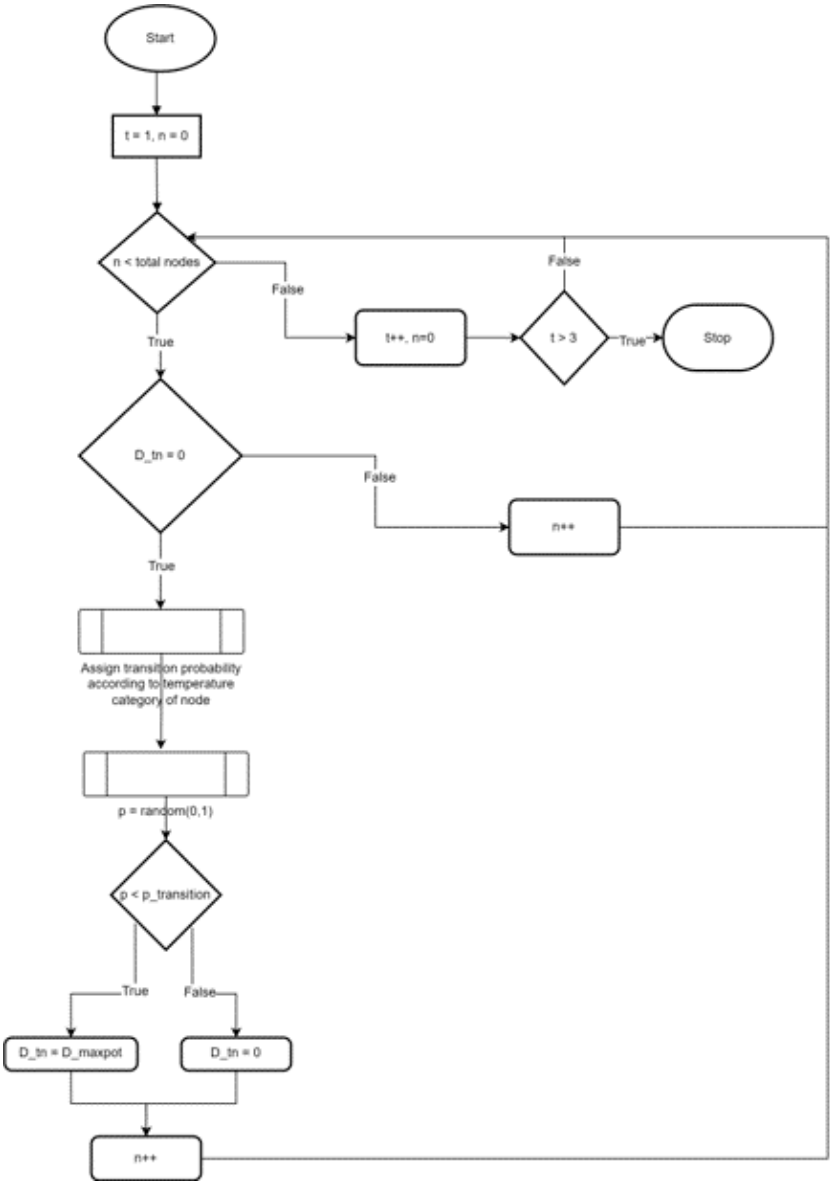


Figure D.1: Flow diagram of the demand scenario generation

E

Details on production capacity distribution

The production capacity range per process is an important variable as it determines the production capacity of the plants in the IPC configuration and thereby the maximum hydrogen demand. The production capacity range per process is derived from the Neuwirth et al. (2022b) German industrial plants database. The representation of the different German industrial processes in the database can be seen in figure E.1. As one can see some processes have only a count of 3, however due to time limitations this database is used to estimate the production capacity distribution. This range of production capacity per process can be seen in figure E.2.

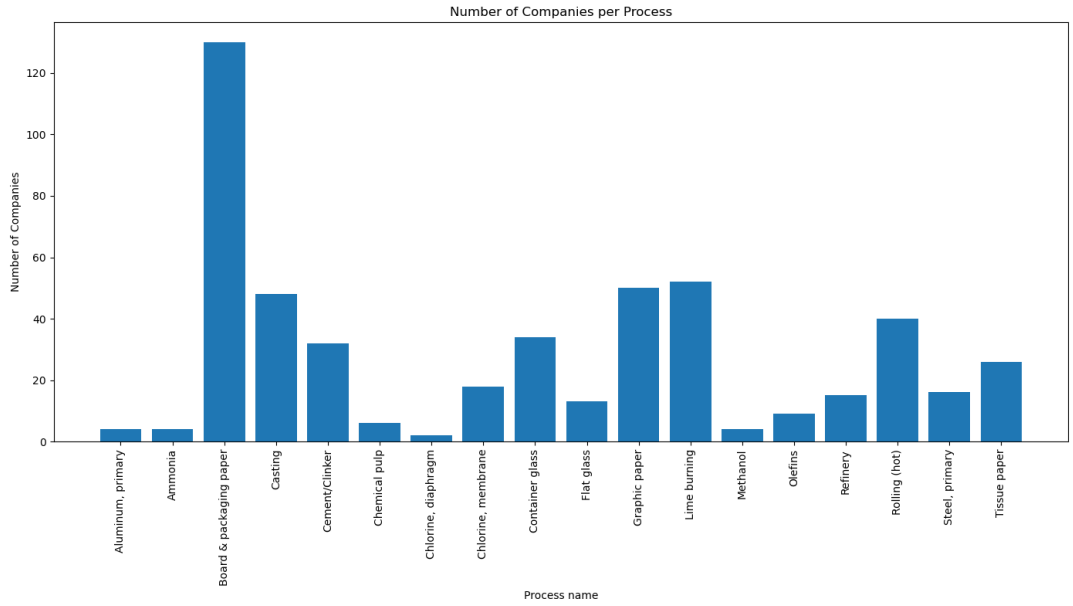


Figure E.1: Representation of the German industrial processes database by counts per process derived from Neuwirth et al. (2022b).

For this thesis the production capacity distribution is chosen to be normally distributed; as it is likely that the capacity of plans is more often average than high or low. In table E.1 the normal distribution parameters that define the production capacity range per process can be found.

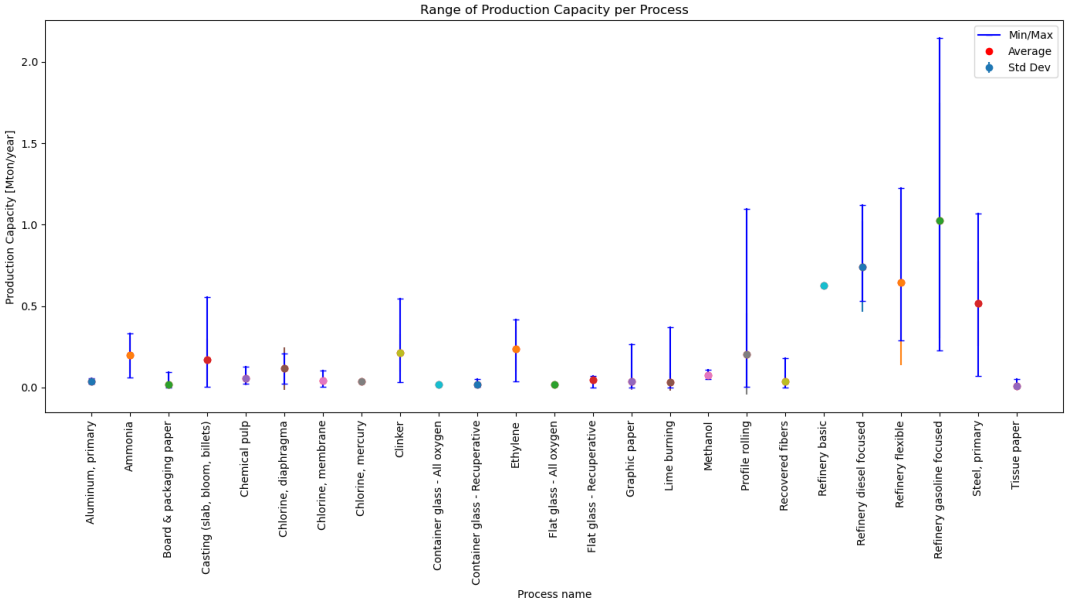


Figure E.2: Range of production capacity of the German industry per process (Neuwirth et al., 2022b).

Table E.1: Production Capacity distribution as derived from the production data from Neuwirth et al. (2022b).

Process	μ [Mton/year]	σ
Aluminum, primary	0.04	0.01
Ammonia	0.20	0.11
Board and packaging paper	0.02	0.02
Casting	0.17	0.12
Cement/Clinker	0.21	0.11
Chemical pulp	0.05	0.04
Chlorine, diaphragm	0.12	0.13
Chlorine, membrane	0.04	0.03
Container glass	0.02	0.01
Flat glass	0.04	0.02
Graphic paper	0.04	0.05
Lime burning	0.03	0.06
Methanol	0.08	0.03
Olefins	0.24	0.14
Refinery	0.85	0.55
Rolling (hot)	0.20	0.25
Steel, primary	0.51	0.26
Tissue paper	0.01	0.01

F

Overview parameters RBPM

G

Overview parameters RBPM

Table G.1: Parameters for the RBPM method

Name	Description
Cost function	This cost function how the network is optimized
IPC parameters	
IPC size	Determines total area of the IPC
Grid size road network	Determines the minimum length of edges between nodes
Number of industrial plants	Number of industrial plants, plants that can transition to hydrogen each decade
Number of extra plants	Number of extra plants that will settle in the IPC given the settlement probability
Number of supply nodes	Number of supply nodes, is assumed to be 1
Location of the supply node(s)	Determines the placement of the supply node, is now assumed to be close to the open water, for safety and port logistical reasons
Cluster coefficient	Fraction determining the amount of plants that belong to the same subsector in a cluster
Demand range	Uncertainty Parameter, that indicates the uncertainty between the calculated potential demand, and the real demand.
Node parameters	
Transition probability (low, medium, high) p_T	Uncertainty parameter, representing the chance of a plant transitioning to hydrogen in a decade that has not yet transitioned. The transition probability uncertainty range has three categories: low, medium, and high. Indicates the likeliness of a process transitioning
Extra node probability	Uncertainty Parameter, representing the business climate; probability that a plant will settle in the IPC.
Production capacity	Constant assigned to a node; indicating the production capacity of the plant. This parameter linearly influences the potential hydrogen demand of a node
Subsector	Subsector that the plant belongs to, plants with the same subsector are economically linked
Process	Process that is executed by the plant
Decarbonization technology	Hydrogen technology that can decarbonize the process
Specific Energy Coefficient (SEC)	The amount of energy needed to produce a ton of produce
Technology readiness level (TRL)	How advanced the technology is from basic idea to commercially available

H

Overview configurations used in experiments

H.1. Plant specifics of configuration proof of concept

Table H.1: Overview of the industrial nodes process specifics, production capacity and hydrogen potential from the example IPC

Terminals	Subsector	Process	Temperature Category	SEC (MWh/t)	TRL (-)	Production Capacity (Mton/year)	Hydrogen potential (ton H2/day)
0	Iron and Steel	Steel, primary	High	1.89	8	0.61	260
1	Iron and Steel	Steel, primary	High	3.20	9	0.86	620
2	Iron and Steel	Steel, primary	High	3.20	9	0.32	230
3	Iron and Steel	Steel, primary	High	3.20	9	0.63	450
4	Non-ferrous metals	Aluminum, primary casting	High	0.60	9	0.01	2
5	Iron and Steel	Steel, primary	High	1.89	8	0.62	260
6	Refineries	Refinery		0.59	9	1.12	150
7	Iron and Steel	Steel, primary	High	1.89	8	1.09	460
8	Iron and Steel	Steel, primary	High	1.89	8	0.55	230
9	Supply						
10	Iron and Steel	Steel, primary	High	3.20	9	0.42	300
11	Refineries	Refinery		0.59	9	1.03	140
12	Iron and Steel	Steel, primary	High	1.89	8	0.30	130

H.2. Visual representation configurations experiment 1

For experiment 1, two clusters Iron and Steel and Paper and Pulp are compared over three layouts, where for each cluster five configurations are generated. Below a visual representation of the configurations is given.

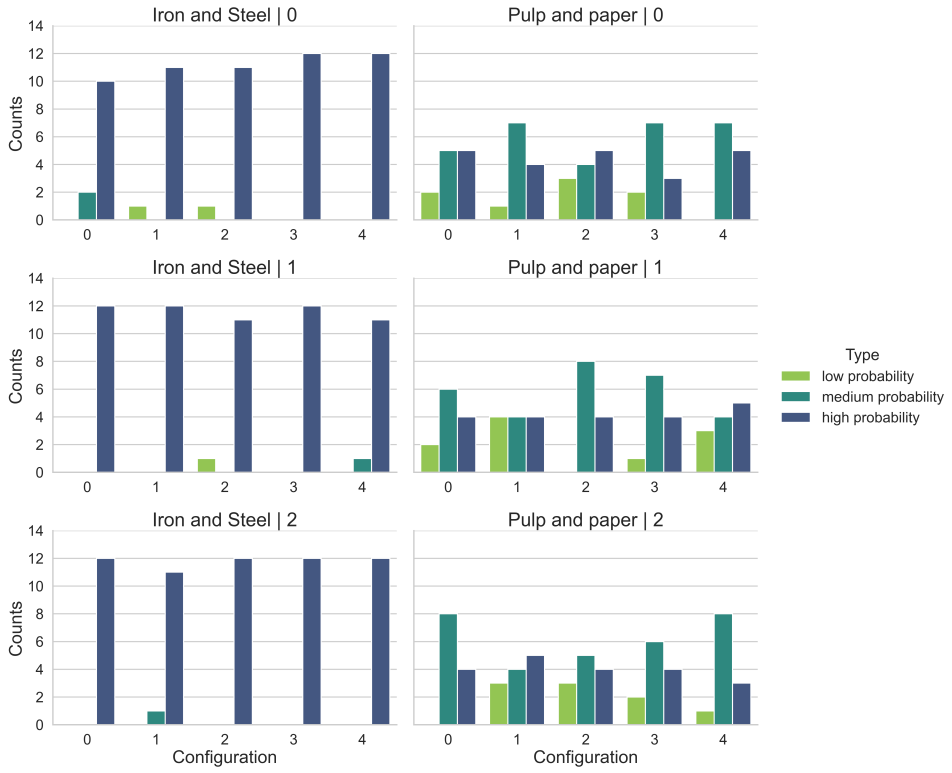


Figure H.1: Overview share of plants with low, medium, or high transition probability in the configurations used for experiment 1

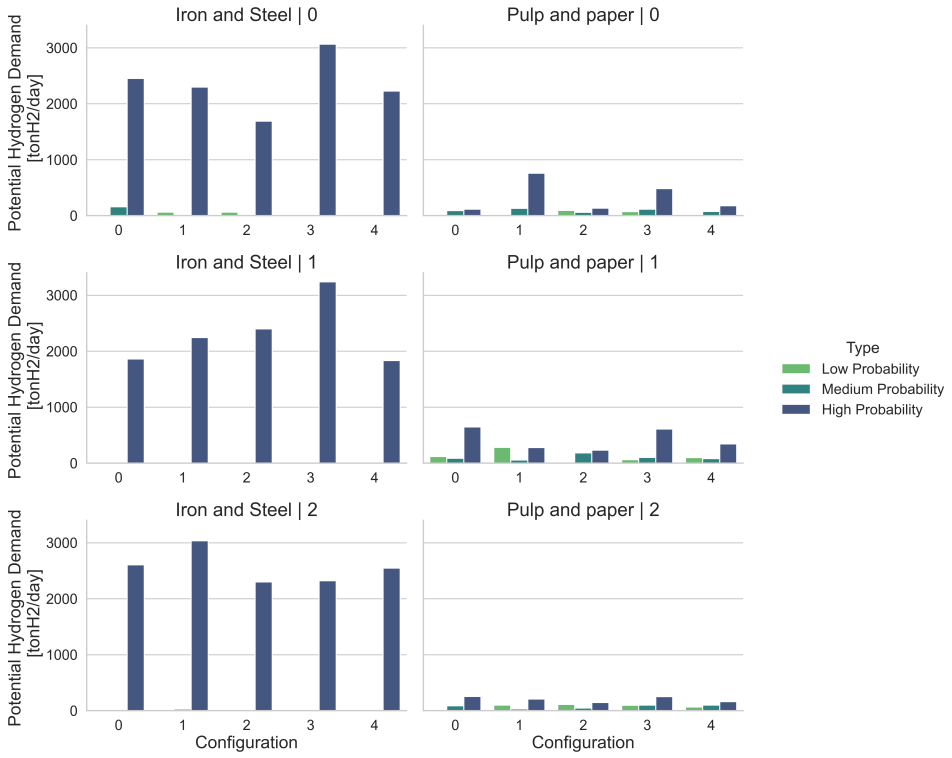


Figure H.2: Overview of potential hydrogen demand categorized in plants having low, medium, or high transition probability in the configurations used for experiment 1

H.3. Visual representation configurations experiment 2

For experiment 2, 8 clusters with each 5 configurations for layout 0 and 1 are used. Below the configurations are depicted per layout by visually representing the distribution in transition probability count and then the corresponding potential hydrogen demand.

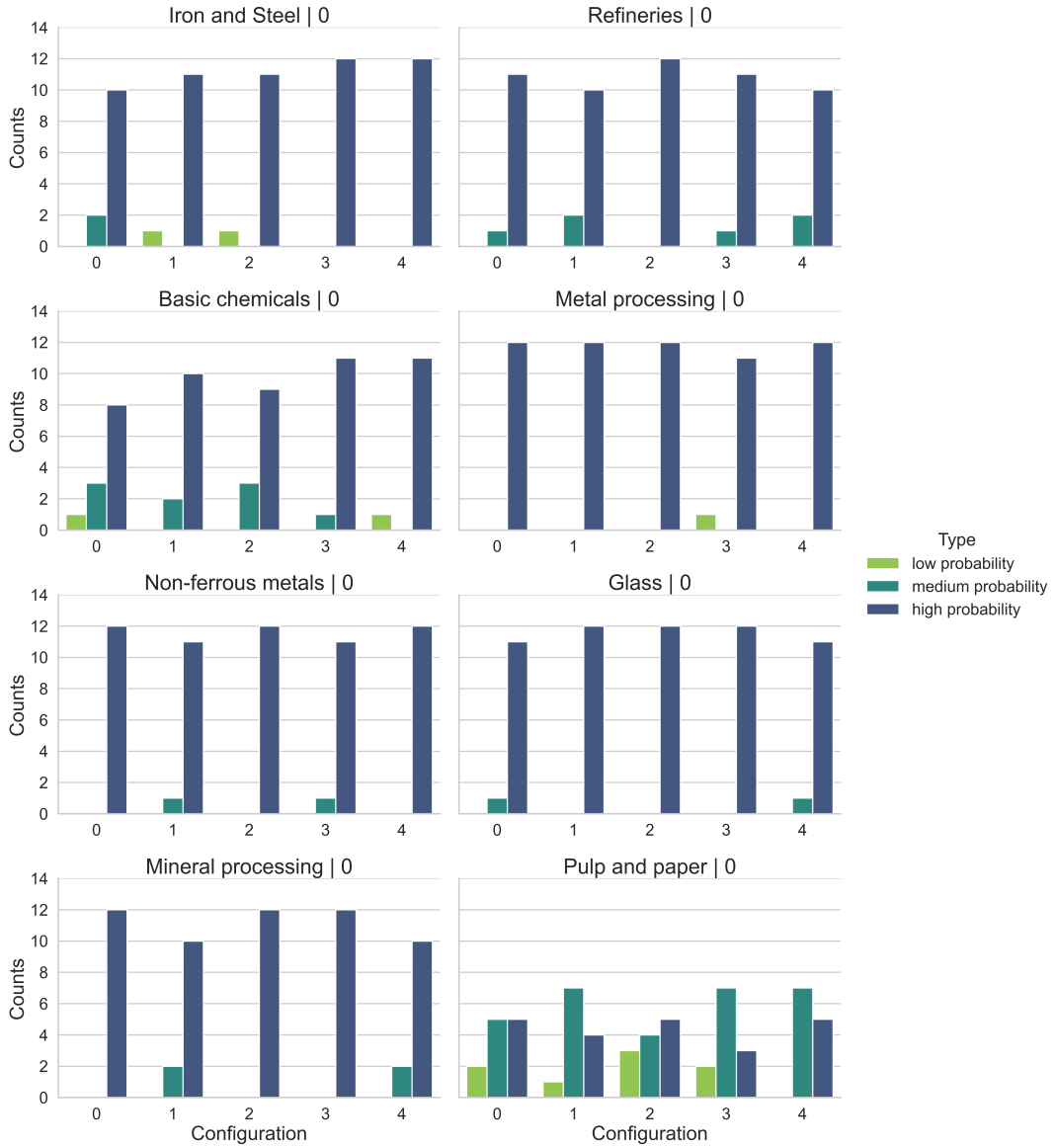


Figure H.3: Overview share of plants with low, medium, or high transition probability in the configurations used for experiment 2, layout 0

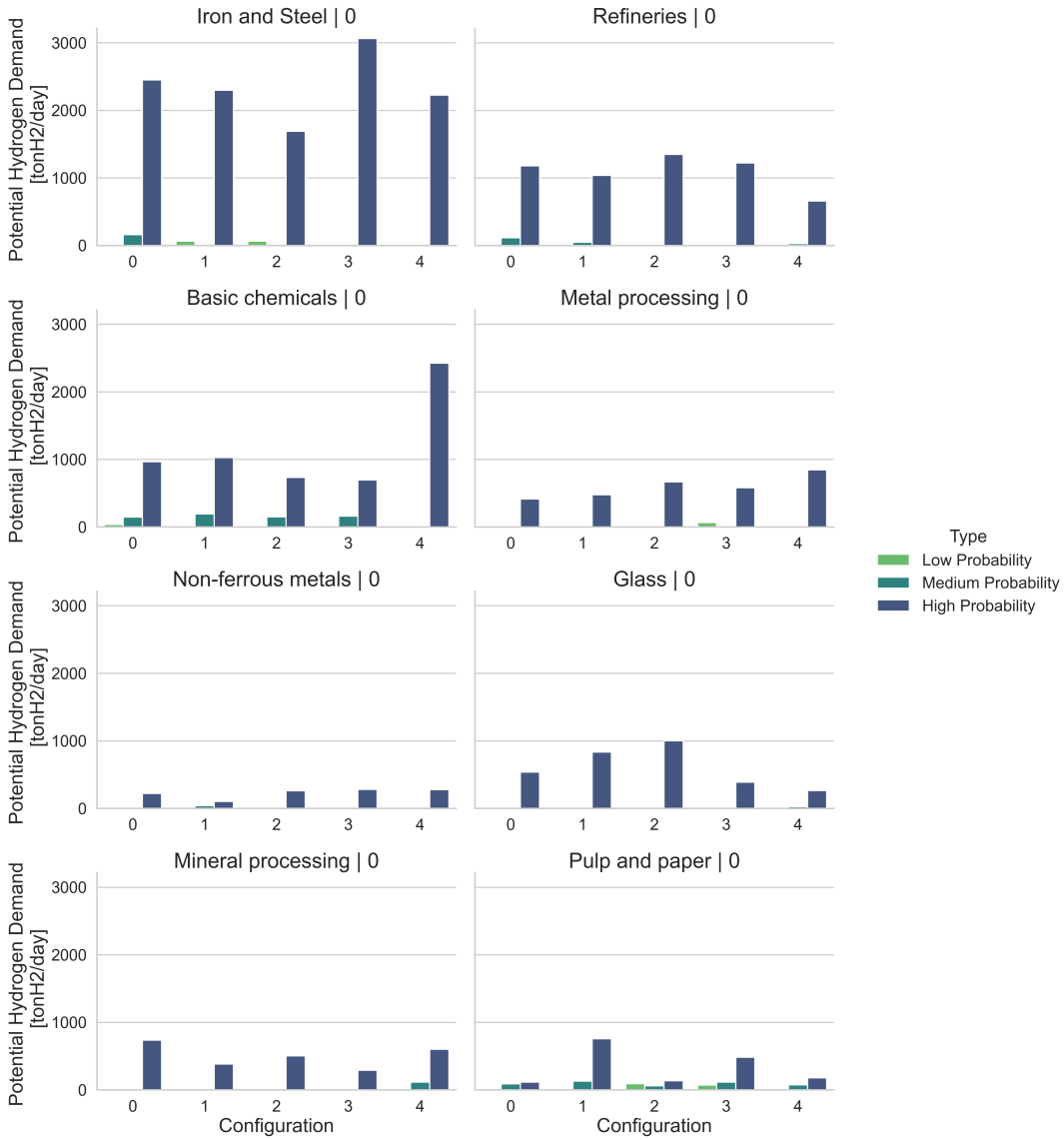


Figure H.4: Overview of potential hydrogen demand categorized in plants having low, medium, or high transition probability in the configurations used for experiment 1, layout 0

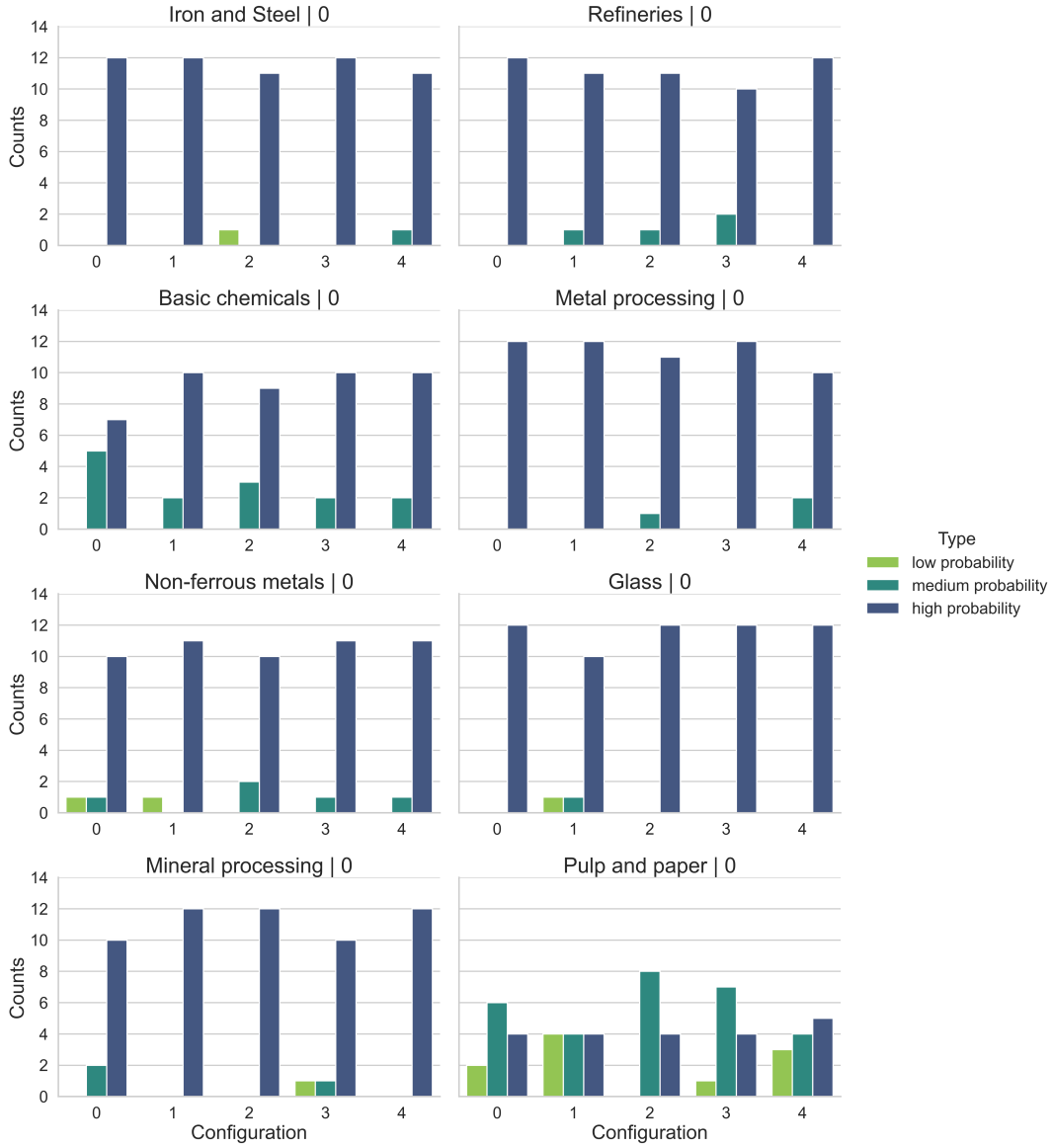


Figure H.5: Overview share of plants with low, medium, or high transition probability in the configurations used for experiment 2, layout 1

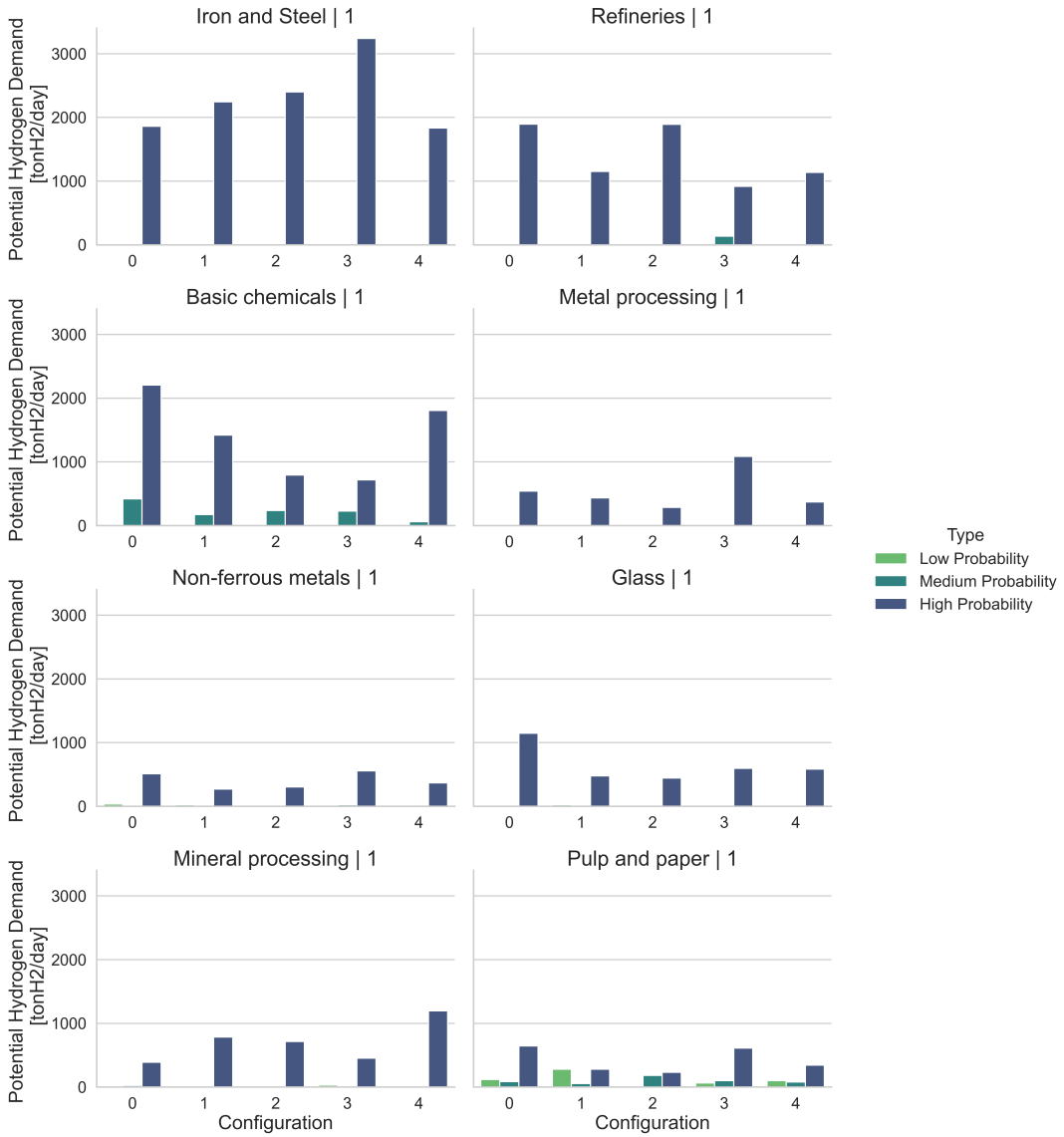


Figure H.6: Overview of potential hydrogen demand categorized in plants having low, medium, or high transition probability in the configurations used for experiment 1, layout 1

H.4. Full IPC Configuration characteristics table

Below an overview of all configurations used and their characteristics regarding placement of the supply node, counts of low, medium, and high probability, the amount of industrial nodes with Technology Readiness Level (TRL) below 6 and the potential future hydrogen demand per transition probability category, and the total potential future hydrogen demand in the whole IPC configuration.

Table H.2: Table indicating the configuration specifics for all experiments, where each configuration is indicated as layout_cluster_configuration.

Configuration	Supply Node	Amount of Plants				Potential Demand [ton H_2/day]			
		Low	Medium	High	TRL $\leq$ 6	Low	Medium	High	Total
0_Basic chemicals_0	6	1	3	8	1	38	146	967	1151
0_Basic chemicals_1	4	0	2	10	0	0	191	1027	1218
0_Basic chemicals_2	4	0	3	9	0	0	150	734	884

Continued on next page

Configuration	Supply Node	Amount of Plants				Potential Demand [ton H_2/day]			
		Low	Medium	High	TRL $\leq$ 6	Low	Medium	High	Total
0_Basic chemicals_3	6	0	1	11	1	0	160	695	856
0_Basic chemicals_4	4	1	0	11	1	16	0	2425	2441
0_Glass_0	4	0	1	11	4	0	13	538	551
0_Glass_1	4	0	0	12	5	0	0	836	836
0_Glass_2	6	0	0	12	4	0	0	1001	1001
0_Glass_3	6	0	0	12	5	0	0	390	390
0_Glass_4	4	0	1	11	7	0	25	265	289
0_Iron and Steel_0	4	0	2	10	1	0	160	2452	2612
0_Iron and Steel_1	6	1	0	11	1	65	0	2297	2362
0_Iron and Steel_2	4	1	0	11	1	64	0	1690	1754
0_Iron and Steel_3	6	0	0	12	2	0	0	3063	3063
0_Iron and Steel_4	6	0	0	12	1	0	0	2227	2227
0_Metal processing_0	6	0	0	12	2	0	0	415	415
0_Metal processing_1	4	0	0	12	4	0	0	479	479
0_Metal processing_2	6	0	0	12	8	0	0	668	668
0_Metal processing_3	4	1	0	11	2	67	0	580	647
0_Metal processing_4	4	0	0	12	3	0	0	847	847
0_Mineral processing_0	4	0	0	12	3	0	0	738	738
0_Mineral processing_1	6	0	2	10	3	0	13	384	397
0_Mineral processing_2	6	0	0	12	6	0	0	503	503
0_Mineral processing_3	4	0	0	12	5	0	0	292	292
0_Mineral processing_4	4	0	2	10	6	0	116	602	718
0_Non-ferrous metals_0	6	0	0	12	4	0	0	221	221
0_Non-ferrous metals_1	4	0	1	11	3	0	42	104	146
0_Non-ferrous metals_2	4	0	0	12	1	0	0	261	261
0_Non-ferrous metals_3	4	0	1	11	0	0	10	283	293
0_Non-ferrous metals_4	4	0	0	12	4	0	0	281	281
0_Pulp and paper_0	4	2	5	5	1	17	91	117	224
0_Pulp and paper_1	4	1	7	4	1	18	131	757	907
0_Pulp and paper_2	6	3	4	5	1	94	60	135	289
0_Pulp and paper_3	6	2	7	3	0	73	114	484	672
0_Pulp and paper_4	6	0	7	5	1	0	75	178	254
0_Refineries_0	4	0	1	11	0	0	114	1178	1292
0_Refineries_1	4	0	2	10	2	0	47	1039	1086
0_Refineries_2	6	0	0	12	1	0	0	1349	1349
0_Refineries_3	6	0	1	11	0	0	14	1223	1237
0_Refineries_4	6	0	2	10	2	0	27	657	684
1_Basic chemicals_0	8	0	5	7	1	0	418	2206	2624
1_Basic chemicals_1	3	0	2	10	0	0	174	1422	1596
1_Basic chemicals_2	3	0	3	9	0	0	236	794	1030
1_Basic chemicals_3	9	0	2	10	1	0	228	716	944
1_Basic chemicals_4	8	0	2	10	0	0	62	1806	1868
1_Glass_0	8	0	0	12	7	0	0	1148	1148
1_Glass_1	10	1	1	10	4	24	10	478	512
1_Glass_2	8	0	0	12	3	0	0	444	444
1_Glass_3	10	0	0	12	6	0	0	596	596
1_Glass_4	9	0	0	12	3	0	0	584	584
1_Iron and Steel_0	9	0	0	12	1	0	0	1861	1861
1_Iron and Steel_1	10	0	0	12	2	0	0	2244	2244
1_Iron and Steel_2	10	1	0	11	2	14	0	2400	2414
1_Iron and Steel_3	9	0	0	12	0	0	0	3240	3240
1_Iron and Steel_4	3	0	1	11	1	0	4	1834	1838

Continued on next page

Configuration	Supply Node	Amount of Plants				Potential Demand [ton H_2/day]			
		Low	Medium	High	TRL $\leq$ 6	Low	Medium	High	Total
1_Metal processing_0	10	0	0	12	7	0	0	540	540
1_Metal processing_1	3	0	0	12	8	0	0	434	434
1_Metal processing_2	8	0	1	11	4	0	6	285	291
1_Metal processing_3	8	0	0	12	5	0	0	1086	1086
1_Metal processing_4	3	0	2	10	5	0	10	372	382
1_Mineral processing_0	9	0	2	10	7	0	26	391	417
1_Mineral processing_1	3	0	0	12	3	0	0	786	786
1_Mineral processing_2	9	0	0	12	4	0	0	715	715
1_Mineral processing_3	9	1	1	10	5	31	20	454	505
1_Mineral processing_4	8	0	0	12	2	0	0	1196	1196
1_Non-ferrous metals_0	9	1	1	10	2	36	7	510	554
1_Non-ferrous metals_1	9	1	0	11	5	25	0	271	296
1_Non-ferrous metals_2	9	0	2	10	3	0	15	303	318
1_Non-ferrous metals_3	9	0	1	11	2	0	24	559	582
1_Non-ferrous metals_4	10	0	1	11	3	0	20	368	388
1_Pulp and paper_0	8	2	6	4	0	119	87	647	853
1_Pulp and paper_1	10	4	4	4	2	282	56	281	618
1_Pulp and paper_2	8	0	8	4	2	0	183	232	415
1_Pulp and paper_3	8	1	7	4	0	64	102	611	778
1_Pulp and paper_4	9	3	4	5	1	100	81	344	525
1_Refineries_0	3	0	0	12	0	0	0	1895	1895
1_Refineries_1	8	0	1	11	2	0	8	1153	1161
1_Refineries_2	10	0	1	11	0	0	9	1892	1901
1_Refineries_3	3	0	2	10	2	0	137	918	1054
1_Refineries_4	3	0	0	12	2	0	0	1137	1137
2_Iron and Steel_0	2	0	0	12	1	0	0	2606	2606
2_Iron and Steel_1	5	0	1	11	3	0	28	3037	3065
2_Iron and Steel_2	5	0	0	12	2	0	0	2300	2300
2_Iron and Steel_3	2	0	0	12	2	0	0	2323	2323
2_Iron and Steel_4	1	0	0	12	3	0	0	2548	2548
2_Pulp and paper_0	1	0	8	4	0	0	86	254	340
2_Pulp and paper_1	2	3	4	5	2	99	33	208	339
2_Pulp and paper_2	5	3	5	4	1	111	46	146	302
2_Pulp and paper_3	1	2	6	4	1	97	100	253	450
2_Pulp and paper_4	1	1	8	3	1	66	99	162	326

I

Additional results

In this section additional figures are presented. They are organized per experiment conducted in Chapter 6.

I.1. Extra figures experiment 1

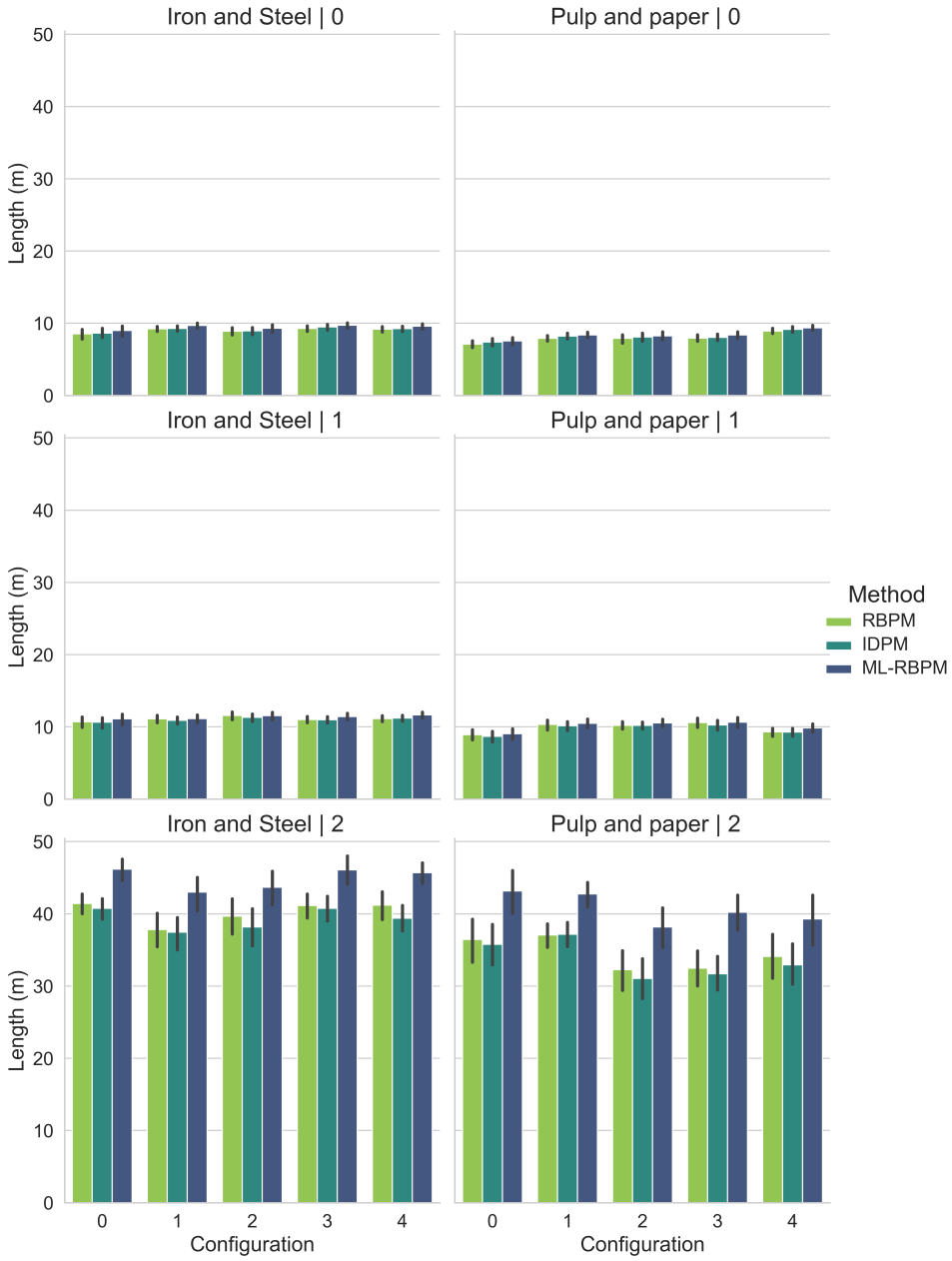


Figure I.1: Comparison of implemented network length for ML-RBPM and RBPM at timestep 3 for clusters Iron and Steel and Paper and Pulp for layout 0, 1, and 2.

I.1.1. Cumulative costs different methods

The three methods are compared per layout by looking at both cumulative costs, and relative costs. In figure I.2, the cumulative costs (CC) average of the five configurations per cluster are depicted. It can be seen that the IDPM performs differently compared with the other two methods. While $CC^{IDPM}(2035)$ is lower compared to the other methods, especially for Iron and Steel clusters, the $CC^{IDPM}(2045)$ and $CC^{IDPM}(2055)$ are often higher.

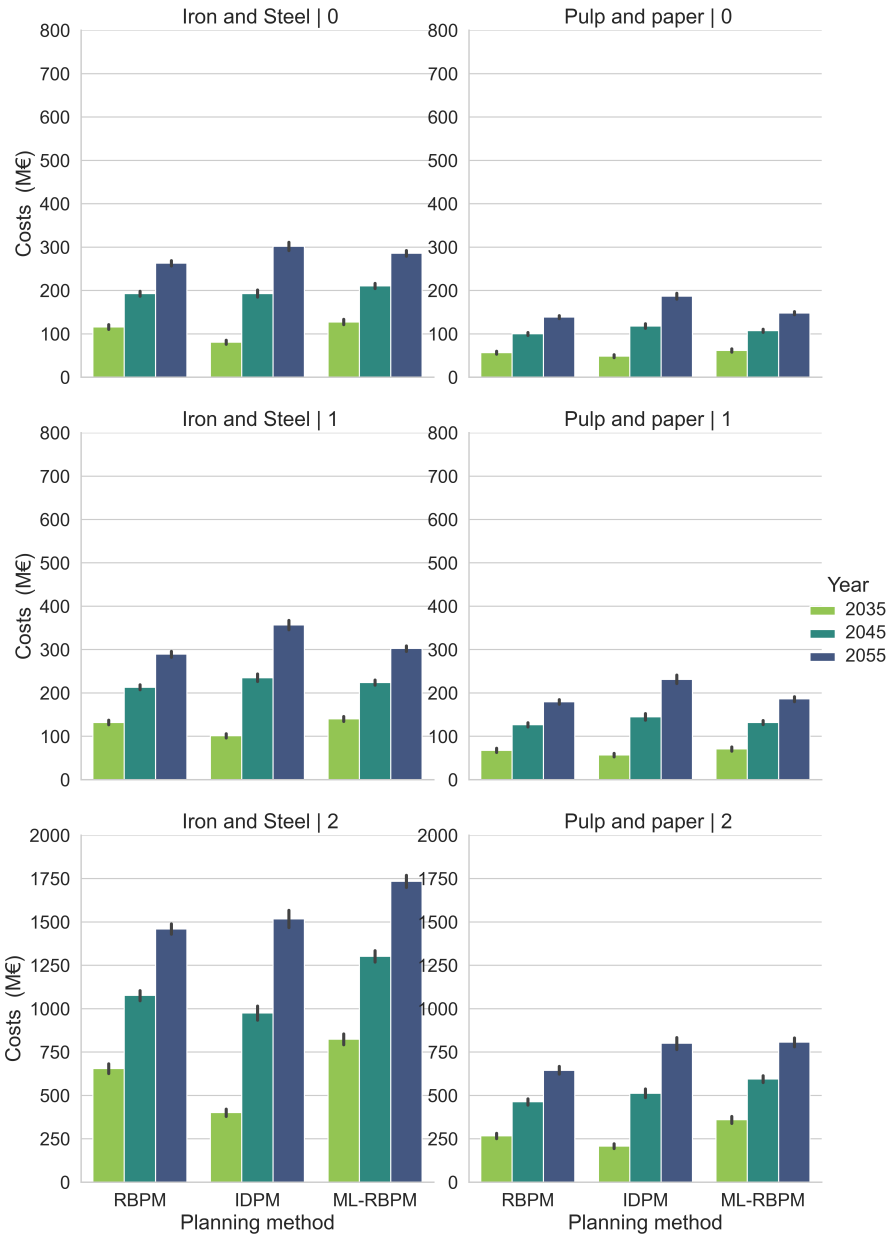


Figure I.2: Cumulative costs comparison Minimum length robust backtracking (ML-RBPM), robust backtracking (RBPM), and immediate demand planning methods (IDPM)

I.1.2. Comparison RN and ML-RN topologies

Below the different robust network generated for RBPM and ML-RBPM are given for the five configurations for layout 2, for the Iron and Steel sector. It can be seen that if the supply node is located on the right side of the water the RN has more similarities with the ML-RN than if the supply node is located to the left side. Specifically, the pipeline network close to the supply node, which has the most capacity, is more similar when the supply location is on the right side. Since high capacity pipelines are costly, the costs for RBPM and ML-RBPM are more similar if the pipelines with high capacity are alike. In other words, taking a detour with low capacity pipelines, is less costly than taking a detour with high capacity pipelines.

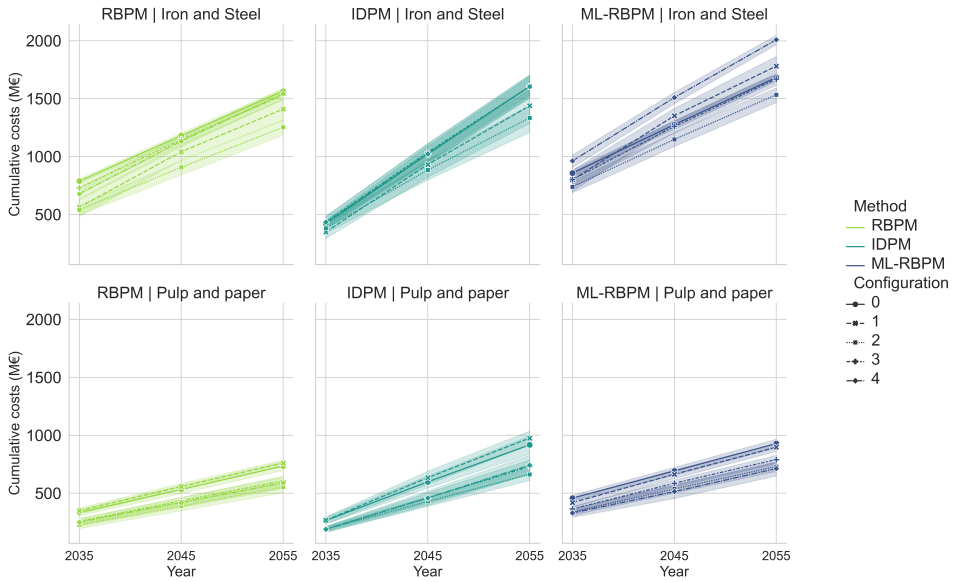


Figure I.3: Cumulative costs of the three methods per configuration and cluster for layout 2

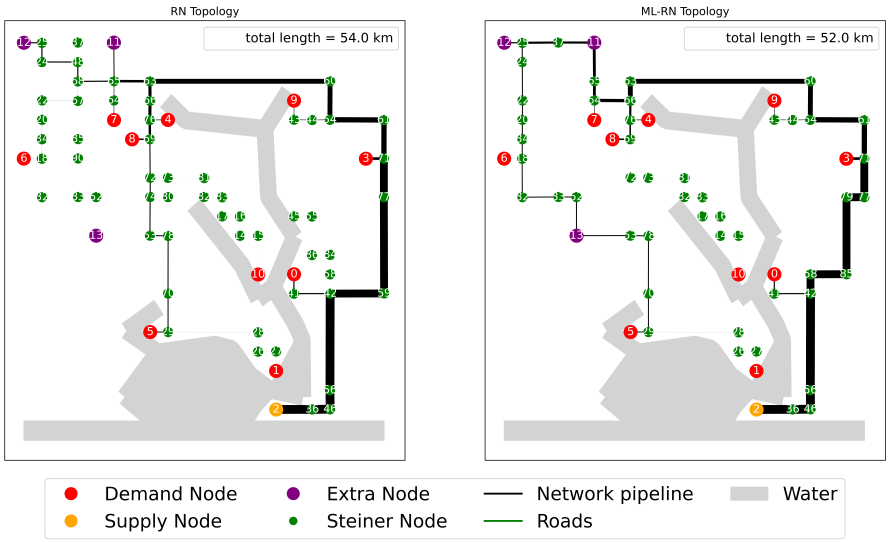


Figure I.4: Topology comparison between ML-RN and RN for configuration 0 Iron and Steel layout 2

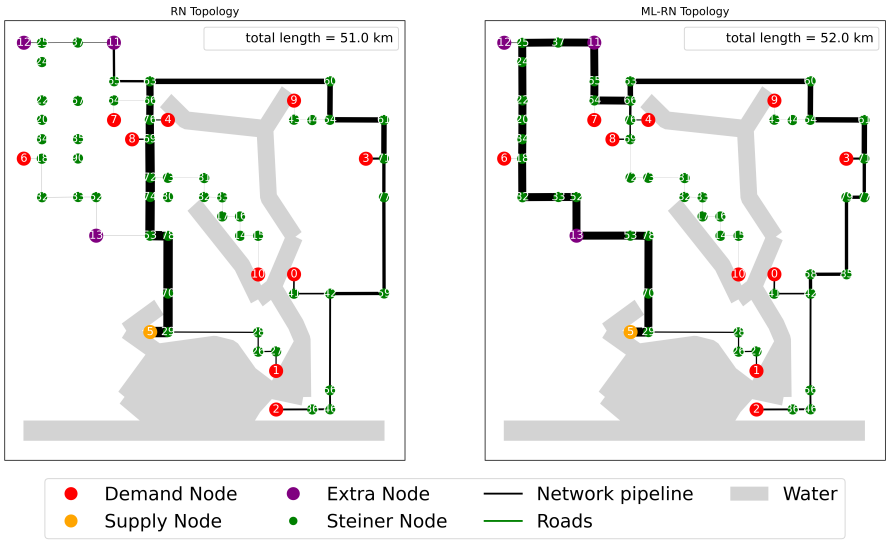


Figure I.5: Topology comparison between ML-RN and RN for configuration 1 Iron and Steel layout 2

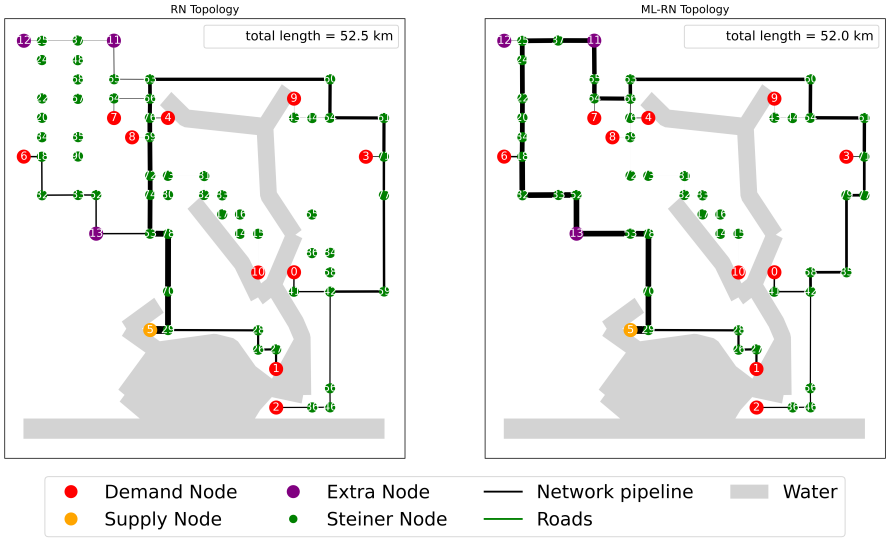


Figure I.6: Topology comparison between ML-RN and RN for configuration 2 Iron and Steel layout 2

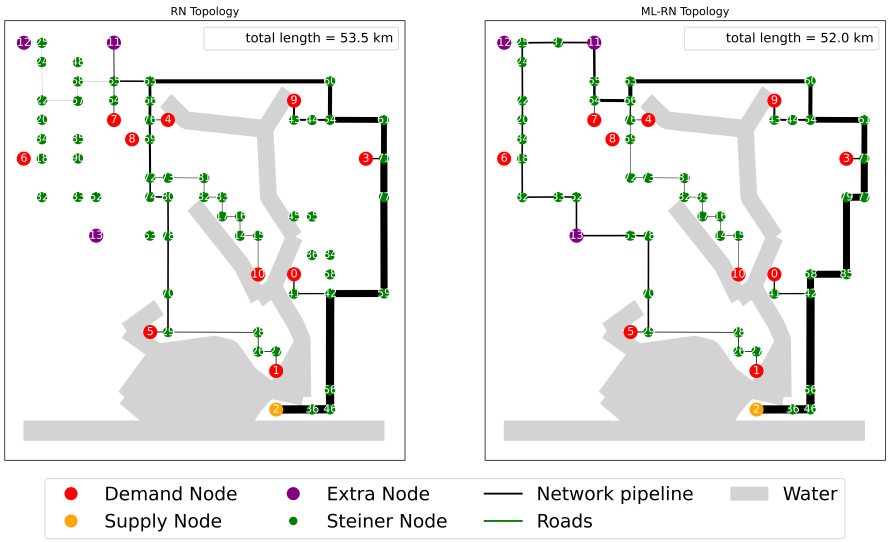


Figure I.7: Topology comparison between ML-RN and RN for configuration 3 Iron and Steel layout 2

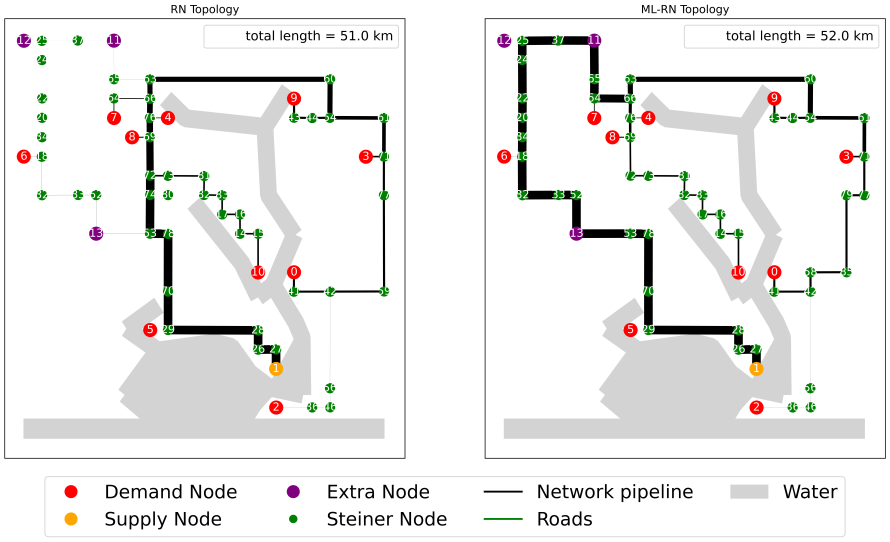


Figure I.8: Topology comparison between ML-RN and RN for configuration 4 Iron and Steel layout 2

I.2. Extra figures Experiment 2

Below, the costs per decade of RBPM for each cluster is depicted in two boxplots. Figure I.9 shows the costs per decade for RBPM C^{RBPM} per cluster. It can be seen that the costs per decade differs per configuration in a cluster. Especially for a cluster such as mineral processing there is a lot of difference between configurations. This can be attributed to the technology readiness level (TRL) of the hydrogen decarbonization technologies of the mineral processing processes.

The TRL of mineral processes is a TRL of below 6, as a consequence these processes cannot transition to hydrogen in 2035. In mineral processing clusters at least 50% are mineral plants, and the other half is picked randomly per configuration. When one configuration has a high share of low transition probability or low TRL plants, this can result in many scenarios where the $C(2035)$ is zero. This can be seen in configuration 2, mineral processing in figure I.9, where more than 25% of the scenarios has zero costs in 2035.

Figure I.9: Cost per decade RBPM for layout 0. For each cluster, the costs per decade are depicted for each configuration

I.3. Extra figures Experiment 3

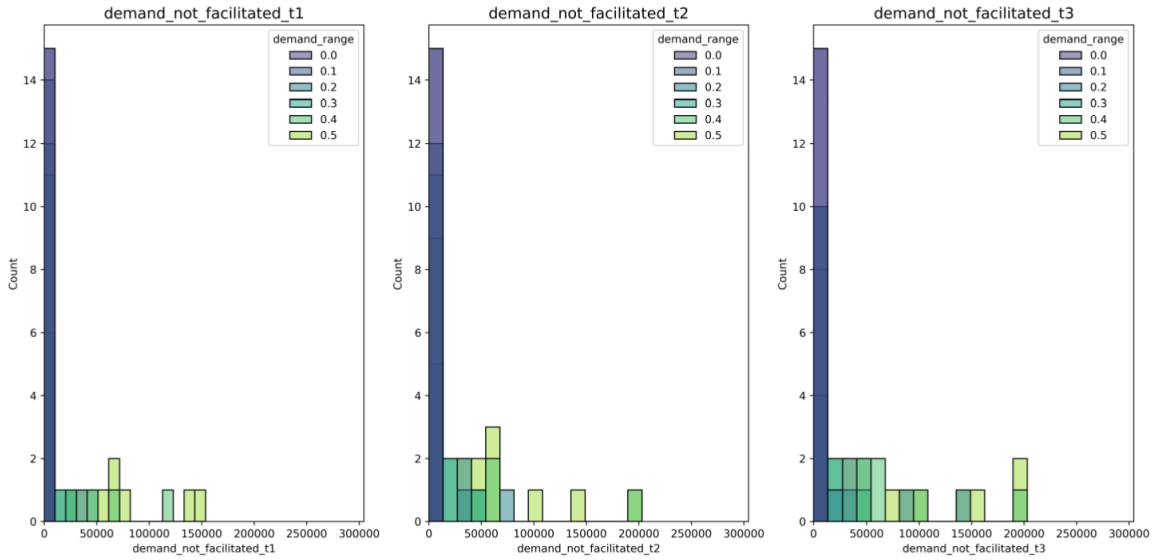


Figure I.10: Histogram indicating of the demand not facilitated per timestep each color represents the demand range related to the analysis scenarios used. The experiment is conducted with higher uncertainty in the analysis scenarios compared to the optimization scenarios. Notably, demand is unmet more frequently in scenarios with high demand uncertainty.