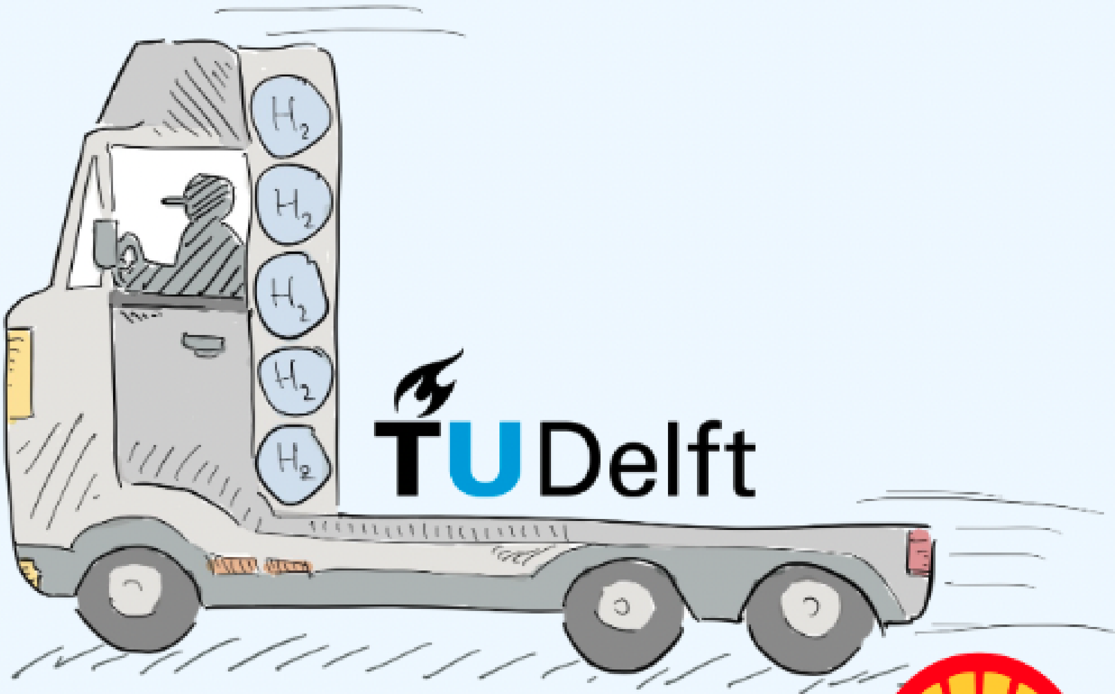
The optimal European hydrogen refueling station network for Heavy Duty Trucks

J.P.J. Siderius

Master Thesis

August 11, 2022

Technische Universiteit Delft



TU Delft Delft University of Technology

Challenge the future

The optimal European hydrogen refueling station network for Heavy Duty Trucks

A master thesis submitted to Delft University of Technology in partial fulfilment of the requirements for the degree of

Master of Science (MSc)
Transport, Infrastructure, and Logistics

Faculty of Civil Engineering (CEG) Faculty of Technology, Policy Management (TPM) Faculty of Mechanical, Maritime and Materials Engineering (3mE)

Technische Universiteit Delft
Delft, The Netherlands

Author:
J.P.J. Siderius
4468481

Supervision: **TU Delft:**
Prof.dr.ir. Z. Lukszo
Dr.ir. P.W. Heijnen
Prof.dr. A.J.M. van Wijk

Shell:
H. Tit

PREFACE

I know that not many of my friends are going to read much of my thesis after this page, so I realize that every word I write here counts.

First of all, I would like to thank my committee. Without Heiko, Petra, Zofia and Ad, this study would not have been at all the work it is today. The critical comments of Zofia and Ad during our meetings have encouraged me to improve my work and set a course to become an engineer. It has definitely given me a better understanding of the energy system and hydrogen's role in it. Thank you very much for that, I know how valuable your time is and it is very much appreciated to get advice from experts like you. My special thanks go to Petra and Heiko. Our (bi-)weekly meetings were of great benefit because they provided me with structure and made me feel like someone was having my back. I very much enjoyed my trips to university (for the first time in my Masters) to meet up with Petra. Next to our discussions on all the plans I had, I really enjoyed our small talk. Small talk which grew quite large at the end. Of course, without Heiko, this study would have never existed to begin with. I am very grateful for the confidence he had in me when he gave me this opportunity. The time I spent at Shell was great, from participating in fun activities to regular days in The Hague: I enjoyed everything. I believe that I have taken full advantage of this opportunity, which even lead to me finding a job.

Secondly, I want to take some time to reflect back on my years as a student. 8 years ago, I wrote my highschool “profielwerkstuk” on “waterstof, de toekomst”. I built my own K'nex car that was powered by a hydrogen fuel cell (which I ordered from the internet). My friend Lucas and I got to visit Shell and the Ecorunner team Delft, both of which I eventually ended up joining. My project was rewarded with an 8. All these years, I thought I had done a stellar job on this PWS. There were some "complicated" calculations in there and we even used proper formatting to write out our formulas. I recently took a look at this report, and it made me realize how much we have grown since then... Without ever taking the time to look back.

I think right now, almost at the end of my educational carreer, is the perfect time for a moment of reflection. Look back, and try to let what has happened since my PWS sink in. Ofcourse first thing that comes to my mind is Maartje. I take it for granted, but starting a relationship at the age of 17 and staying together for 8.5 years is really something else. On the other hand, it feels like we're just getting started. There are some friends who have been around even a little longer. From 0294, to the NMG, all close friends who never lost sight of each other. During my college years, I had the pleasure to add a group of American friends, 223, a club of 20 lads from the Hague, a study board, the Maison, and last but not least: the Matrici. My TPM study group, which I'm sure if it wasn't there, no one would have gotten their degree. This one's for you!

Lastly, I want to thank my family: Jan, Annemiek, Flo, Feije, Diederik, and Diederick. Even after I left my home in Loenen, I shared great times with you. With Died, Died and Flo in Delft and with Feije at Ajax or in Groningen. Not to leave out our annual family vacations. I think little has ever been lacking in our lives: we have all been happy, surrounded by friends, and doing what suits us best. A special thanks for that goes to my parents. I am grateful for the way they raised us and for the time they devoted to us. You made me who I am today.

Ps. don't be intimidated by the 100 pages, there are many nice figures in the report... Enjoy!

This page intentionally left blank.

Executive summary

Heavy-duty trucks make up a small percentage of the European fleet, but emit a large proportion of road freight transport emissions. Electrification for this sector seems difficult, since caving in on range and refueling time has significant impacts on total costs of operations. An emission-free alternative fuel with suitable characteristics is found in hydrogen. Hydrogen is expected to emerge as a storage medium for energy from renewable sources. The need for storage is driven by variable electricity generation, which calls for a more flexible energy system. This stored hydrogen is best used in sectors that are difficult to electrify, such as heavy-duty trucking. One company active in the hydrogen market is Shell. Shell supports the creation of regional hydrogen clusters. These clusters will consist of hydrogen refueling stations combined with local truck demand. Over time, the intention is that these clusters should be linked together to create a full-coverage European hydrogen network.

Shell's network plans to date have been based primarily on diesel data from their own stations. Modeling based on diesel refueling data reflects realistic behaviour as it is based on historical data. However, it also brings with it some disadvantages as (1) diesel refueling behaviour has specific characteristics and (2) Shell diesel data is not complete and specific. This introduces a set of drawbacks that needs to be examined with a scientific approach. Scientific studies seem to be lacking on this topic in level of aggregation and specification for heavy-duty trucks. Hence, the goal of this thesis is to investigate the optimal topology of a European network of hydrogen refueling stations (HRS) for heavy-duty trucks.

To arrive at an optimal topology, performance measures for a network of refueling stations are established. Other studies are consulted and this leads to the following set of criteria; (1) the utilization rate of the HRS; (2) the weighted average distance (travel time) to an HRS; (3) maximization of total vehicle kilometers covered; and (4) minimization of CAPEX (costs for building and supplying stations). This collection of criteria is indicative of the search for balance between costs for supply and demand. Supply is represented by operator costs (CAPEX and utilization rates) and demand is represented by costs of coverage (weighted distances and vehicle kilometers covered). The methodology created for this study sets out to optimize this balance.

To enable optimization, a basis for modeling is established. Truck traffic is modeled over a highway network using a freight demand model. Synthetic traffic flows across Europe are created and recalculated to represent truck vehicle kilometers on a point of highway. Based on this, the geographically dispersed demand for hydrogen can be determined. This allows measurement and optimization of three performance criteria, but not fully for CAPEX. Minimization of CAPEX for this study is determined by the number of stations and costs of supply. The latter cannot yet be accounted for. To do so, only supply by an assumed pipeline model (a hydrogen backbone) is considered. This is a form of delineation that also fills another gap in the literature: a lack of consideration of a backbone when determining the location of stations.

This modeling basis omits the Shell diesel data biases and covers the gaps in the literature. It is accompanied by a system description, which clearly defines the boundaries of this study and results in input for the model. The hydrogen supply chain is analyzed and a policy analysis is performed. The latter results in two constraints that determine station locations by regulation. Both constraints are currently being negotiated under the Alternative Fuels and Infrastructure Regulation (AFIR). Since they are not yet final, they are analyzed for the purpose of recommendations, but not taken as the basis for this study.

To model a network of hydrogen refueling stations based on the balance of demand and supply, the methodology prescribes a two-step optimization:

- First, a station location model is used to create an optimal demand-based topology. A weighted k-means algorithm is used to cluster the highway data points to a predetermined number of stations. Weighted k-means clustering ensures optimization of the demand performance measures as the algorithm minimizes the weighted distance of the datapoints to the stations while covering 100% of vehicle kilometers. This result is further optimized in terms of CAPEX (built stations) and utilization by excluding stations below a demand threshold. The threshold is determined by an optimal balance between coverage, number of stations and the utilization rates. A demand-optimal topology results which provides initial insights into the market.

- Second, to further optimize the network in terms of CAPEX, supply is included. By assuming a hydrogen backbone, and placing that in the same plain as the demand-based network, use of an optimization tool is made possible. This Optimal Network Layout Tool connects each demand-based station (per pipe) to the assumed backbone in optimal fashion. Hence, an optimal network system is created. This releases output for further analysis of supply costs in the form of required capacity and lengths of pipelines.

- Finally, the demand-based topology is re-optimized with influence of supply. A balance between the costs of supply and demand is achieved, completing the two-step optimization.

This methodology seems to steer toward a single design that would be the end result. However, an approach is adopted of searching for this design while collecting key outcomes along the way. These key findings lead to recommendations that are the real value, since in reality no single design is perfect for network planning.

- First, analysis on the AFIR proposed station sites leads to the recommendation for Shell to only lobby for subsidies on stations in the peripheral areas of Europe if necessary. Only those sites will be underutilized and further support for the regulation will help create a pan-European network.

The other recommendations are intended for the network designed by Shell. It is recommended that an integral full-scale network is designed today based on 2030, 2040, and 2050 expectations.

- The early-phase station locations are advised to be determined based on demand as a result of the 2030 demand-based network. Demand coverage is important for initial market traction and the backbone in 2030 does not yet enable high coverage of supply. For this network, an expected demand threshold of 1 ton per day for stations to be built avoids severe underutilization. In addition, this greatly reduces CAPEX by building far fewer stations while maintaining high coverage.
- On the supply side, average station size appears to allow for significant cost reductions. Since the initial stations are based on demand, to balance, the average station size in 2040 is modeled at 6 tons per day. CAPEX is reduced both in terms of levelized transportation costs and the number of stations to be built. In light of the aforementioned supply performance criteria, this scores well, but it will not increase coverage or further reduce driving distances. For this reason, it is recommended that the average station size remains the same towards 2050.
- When demand increases but the average size of stations remains the same, more stations should be built. By 2050, the backbone will have reached a form that allows for the supply of much of the HRS. Once a backbone is in place, the degree of interconnection to the backbone appears to be a key factor for further cost optimization. A high degree of interconnection leads to relatively low supply costs, since pipeline capacity can be shared. As many new stations will be built in the decade to 2050, the opportunities for interconnection to the backbone increase. Station locations should be determined with this front of mind.

Altogether, this study proposes a method for designing a network of hydrogen refueling stations: first by demand, second by growth in station size, and third by interconnectivity to a hydrogen backbone. This will allow for market traction in the first place and cost savings over time. It shows a fluctuating balance of supply and demand. For this thesis, this balance is sought by making many assumptions and disregarding many factors. For example, the emergence of multi-modal stations or shared use with private vehicles could further influence this balance in terms of supply costs. On the other hand, higher market penetration of hydrogen in heavy-duty trucking could further increase the influence of demand. Nonetheless, the results are indicative of optimal station locations and add value when interpreted correctly.

For Shell, it is therefore important that these results are properly incorporated into network planning. In this light, it is recommended that the resulting maps are treated as layers in a larger design. The study assumes that an HRS can be built anywhere. Therefore, first layers to be added could be realistic new station locations (greenfields) and the current locations of Shell stations that can be retrofitted or expanded. Placing these layers over each other reveals the real possibilities for optimal station locations. Further studies are recommended to add more layers to the puzzle. The most obvious opportunity seems to lie in the specification and diversification of supply. Specific hydrogen supply sources are not considered in this study and different distribution types may fit each station. Regardless, designing an integrated network for hydrogen refueling stations starts today, begins in 2030, grows in 2040, and connects in 2050.

Contents

List of Figures

LIST OF TABLES

Nomenclature

AFIR	Alternative Fuels Infrastructure Regulation
AFP	Alternative Fuels Powetrain
CAPEX	Capital Expenditures
ConFLP	Continuous Flow Location Problem
DSO	Distribution System Operator
EU	European Union
FC-HDT	Fuel Cell Heavy-Duty Truck
FRLM	Flow Refueling Location Model
HBB	Hydrogen Backbone
HDT	Heavy Duty Truck
HFCV	Hydrogen Fuel Cell Vehicle
HRS	Hydrogen Refueling Station
IPCC	Intergovernmental Panel on Climate Change
IEA	International Energy Agency
NC-FRLM	Node Capacitated Flow Refueling Location Model
OD	Origin Destination
OPEX	Operating Expenditures
PEM	Polymer Electrolyte Membrane
PSA	Pressure Swing Adsorption
UN	United Nations
TSO	Transmission System Operator

1

INTRODUCTION

Global climate change has already had noticeable impacts on the environment and effects that scientists had predicted in the past are now occurring. Rising sea levels, melting glaciers, frequent and intense droughts, storms, heat waves, and warming oceans: the effects of human-caused global warming are irreversible in the time scale of people living today, and will worsen in the coming decades [17]. Scientists are confident that global temperatures will continue to rise in these decades, mainly due to greenhouse gases produced by human activities. At the global scale, the main greenhouse gases emitted by human activities are carbon dioxide (CO2), methane (CH4), nitrous oxide (N2O), and fluorinated gases (F-gases) [18]. Of these gases, CO2 is the most notorious, and for good reason: carbon dioxide alone accounts for 76% of total greenhouse gas emissions. The Intergovernmental Panel on Climate Change (IPCC) forecasts that these emissions will make temperature rise with 1.5 to 5.5 degrees Celsius over the next century [19]. The magnitude of impacts on individual regions will vary locally and over time. On a continental scale, Europe will be affected differently across the landmass; southern and central Europe will experience more frequent heat waves leading to forest fires and droughts; the Mediterranean will become drier, making it more vulnerable to droughts and forest fires; northern Europe will become significantly wetter, and winter floods could become very common [20].

To combat these effects of climate change, European governments are trying to implement methods to reduce greenhouse gas emissions. By implementing the European Green Deal, the continent targets to become the world's first climate-neutral continent by 2050. Recent years have shown that their efforts, in combination with those of the other UN, have not been sufficiently successful. Even if all the measures prior to COP26 (a 2021 UN climate conference) would be implemented on time, an analysis by WEO-2021 predicts that the world was headed for a 2.1°C warming by the end of the century [21]. The goals of the Paris Agreement would not have been met, with all the consequences that this entails. However, in light of COP26, more countries have revised their ambitions upward. Major economies have pledged to achieve net zero emissions. An analysis by the IEA shows that if the new commitments of COP26 are fully and timely implemented, they will be sufficient to limit the rise in global temperature to 1.8°C in the year 2100 [22]. A milestone, as it is the first time governments have set targets ambitious enough to limit global warming to less than 2°C.

To achieve this ambitious plan, it is crucial to implement solutions that have impacts across the many sectors that cause global warming. So far, Europe has achieved the most emission reductions by changing the way it generates electricity. The burning of coal has been reduced and the use of hydro, wind and solar power sources has increased significantly. Eurostat figures show that in 2019, 28% of gross electricity consumption in the EU was generated by wind and water [23]. This shift implies a higher share of variable renewable energy in electricity generation, making it difficult to maintain a constant electrical frequency. This balance is important as multiple frequencies cannot operate side by side without damaging equipment. To deal with this imbalance, energy storage could provide flexibility and thus support renewable energy integration in the energy system. For decades, energy storage has already contributed to the system by storing electricity when demand is low and releasing it when demand is high, usually on a daily cycle. However, new generation patterns are changing the requirements in the sense that not only short-term storage, but also seasonal storage must be possible. These changes bring along difficulties, but also new opportunities to utilise energy storage, as the electricity sector can be linked to efforts to decarbonize industry and transportation.

1.1. HYDROGEN

The European Commission identifies storage in the form of hydrogen as a means to resolve the imbalance created by renewable electricity generation. This is mainly due to its ability to store large amounts of energy on timescales of weeks and months. Figure 1.1 shows the storage capacity of hydrogen, compared to different energy storage options.

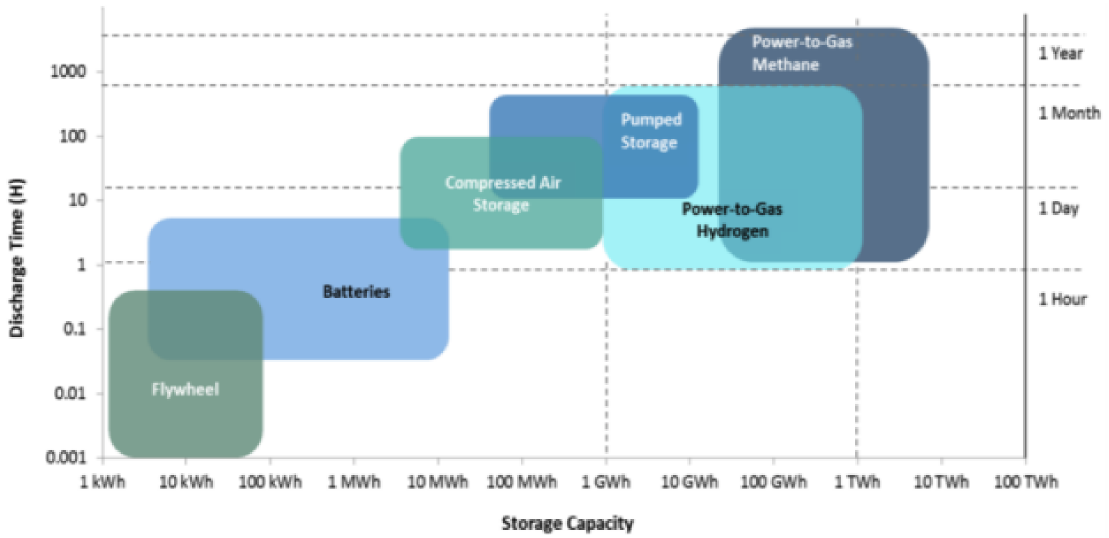


Figure 1.1: Electricity storage capacity and discharge time of different energy storage mechanisms [1]

As the figure shows, large amounts of energy can be stored in the form of hydrogen, significantly increasing the flexibility of the electricity system. The hydrogen can be made by electrolysis fed by surplus renewable electricity. In REPowerEU, the EU announced its hydrogen strategy with the ambitious goal of using 20 MTon of hydrogen by 2030, all renewable electricity-fed [24]. Currently, this market is barely a fraction of this goal and the regulation would be an attempt to make renewable and low-carbon hydrogen cost competitive with fossil fuel-based hydrogen. Effectively, they could produce up to 10 million tonnes of renewable (green) hydrogen each year by 2030. The molecules can then be stored for long periods of time and on a large scale, making it the most interesting candidate for the enablement of seasonal balancing of renewable energy. Added to this, is the advantage that the hydrogen, once generated, can be transported by ship, truck or pipeline at reasonable cost [25] enabling potential decarbonization of other sectors. This characteristic allows for international transport, which means that regions with favorable conditions for generating renewable energy can become exporters of large quantities of hydrogen further reducing costs. The Hydrogen Council concluded that this feature of the energy carrier is a large factor of production costs declining faster than previously thought [4].

Taking into account hydrogen's ability for long-term storage of renewable energy, the question arises as to where this energy is most useful. To answer this question, one could consider which fossil fuel could be best replaced by hydrogen. This replacement would both be in terms of current costs and supplying infrastructure of the fuel in question. In this light, the transportation sector emerges as a field where both cost-effectiveness (due to relatively high gas prices) and a large infrastructure network are present.

Currently, the transportation sector is showing promising changes in the area of driving on electricity. Because it is a zero-emission option that already covers a large network, there seems to be little room for hydrogen. However, this is currently only the case for smaller cars, vans and short-haul trucks. There is a reason why the impact of battery transport on long-haul trucks is slow to take off: the sector asks for flexibility and high productivity, which is difficult to provide with current technology of batteries in terms of range and charging speed [2]. Fast charging is technically feasible, but it also involves high operating costs, resulting in high total cost of ownership [26]. Hydrogen is therefore emerging as an attractive solution for the trucking market in specific. Analysis of fuel cell trucks suggests that this technology is the lowest-cost way to decarbonise both the medium- and heavy-duty segments [2]. Fuel cell heavy-duty trucks (FC-HDTs) and coaches are estimated to achieve cost parity by 2030 as can be seen in the figure 1.2 below.

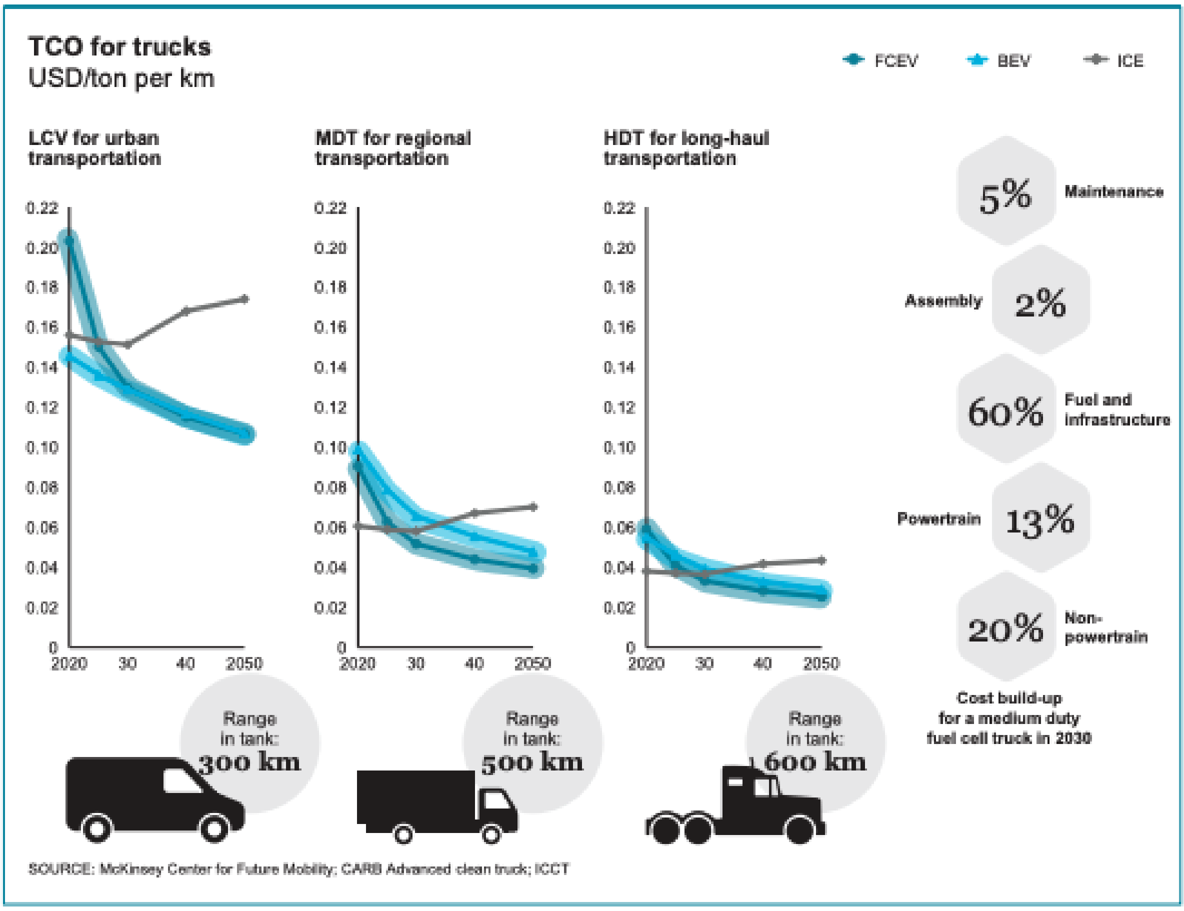


Figure 1.2: Cost-effectiveness of heavy-duty trucks [2]

1.2. HYDROGEN IN TRUCKING

Of the 6% of total EU emissions that is caused by heavy-duty transport, a large part is caused by the trucking sector. The long-haul trucks make up a small 12% of the European fleet but emit 41% of road freight emissions due to heavy weight, high payload and long distances travelled [27]. The path to zero emissions for this sector is a difficult one, since caving in on range and refuelling time has significant impacts on total costs of operations. Therefore, an alternative is found in hydrogen that is able to achieve a large vehicle range (up to 800km even for the heaviest loads), with a refueling time similar to current diesel trucks (less than 15 minutes) [28]. Other options with similar fuel characteristics include biofuels or the use of hydrogen for internal combustion. While these alternatives are well placed to reduce CO2 emissions from freight transport in the short term, they will have little effect on other greenhouse gasses associated with internal combustion. It therefore seems inevitable that in the long term hydrogen fuel cells will be incorporated into long-haul trucks. One company looking to jump on this gap is Shell, as will be explained in the next section.

1.3. THE ROLE OF SHELL

To accelerate the advent of cheap and abundant green hydrogen, the European Commission has launched a new hydrogen research partnership in the European Clean Hydrogen Alliance [29]. A major player in the energy sector that has joined this alliance is Shell. Shell is signaling its intention to play a major role in hydrogen production and contribute to Europe's goal of achieving climate neutrality. For example, they recently commissioned the largest PEM green hydrogen electrolyser in Europe [30]. Shell sees great potential for the use of hydrogen in a range of sectors, they categorize their own interest in three sectors: heavy-duty transport, light duty transport, and the industry [31]. For heavy-duty transport, Shell identifies hydrogen to have the potential to be an important, safe, low-carbon transport fuel. Their focus lies on transport modes such as trucks, buses and shipping. Their involvement is strengthened by their role in H2Accelerate, a mobility consortium that aims to create the conditions for a large-scale rollout of hydrogen infrastructure for trucks in Europe over the next decade. In this consortium, they work together with Daimler AG, IVECO, OMV and Volvo Group, TotalEnergies, and Linde to prove the viability of hydrogen for heavy-duty in Europe. The major advantage of such a consortium is its ability to address the infrastructural chicken-and-egg problem. The problem being that refueling infrastructure is not rolled out when there are no trucks, but there will be no trucks on the road without refueling infrastructure. Currently, the consortium is tackling this by jointly scaling up in two phases; (1) rollout of first stations and trucks and (2) Europe-wide coverage [32]:

Table 1.1: Phases of H2Acceleration rollout

Phase 1	**Phase 2**
Rollout of firs stations and trucks	Europe-wide coverage
100's of trucks	Second half of 2020s: Achieve volume manufacture '000's per year
>20 high capacity stations	Rapidly reaching >10,000 trucks
Proving high capacity station concepts	Europe wide coverage of major corridors
Selective locations/clusters	High capacity/reliability stations

The plan spans a decade, as it aims to reach its final stage in 2030. To deal with infrastructural scaling problems, they are aiming for separate groups of customers willing to commit to hydrogen freight transport at this early stage. These customers will operate in regional clusters, but over the course of the decade, these clusters should be linked together to create a full-coverage European hydrogen network.

1.4. A European Hydrogen Refuelling Network

To achieve a pan-European hydrogen refuelling network, much remains to be done. On Shell's timeline, we are currently in phase 1. The character of this phase is that of "learning by deployment" [33]. The combination of research and development, and early deployment acts as a tool to define the broader market introduction phase. As such, there is a significant amount of learning to be done. Specific aspects of what still needs to be learned are described as follows [33]:

- Understanding customer attitudes to the hydrogen trucks and the implications for future roll-out (network design, technology choices etc.).
- Testing the economics of both liquid hydrogen and gaseous hydrogen
- Gaining new knowledge about the longevity and failure rates for the components within the supply chain regarding the vehicles and the hydrogen refuelling stations (to develop a view towards a more robust systems in the post 2025 roll-out).
- Deploying and testing of hydrogen refuelling stations with a high capacity of 1 to 2 tonnes per day of refuelling capacity and are capable of truck refuelling in 10-15 minutes.
- Developing new designs (and components) for stations with a capacity in excess of 5 tonnes per day, with faster filling than for the 1-2 tonne/day station.
- Finding the strategies which can allow large scale renewable hydrogen production to grow along with a fleet of zero emission trucks.

As can be concluded from these aspects, scientific research will accompany the roll-out plans of Shell. As such, Shell is looking to find the optimal network design to locate their refueling stations as is indicated in the first aspect of understanding customer attitudes. The optimal design of the refueling stations network depends on the costs of distance to supply and demand. The distance to demand is represented by the distance to potential customers (trucks). The supply of refueling stations can take place via pipeline, or the hydrogen can be transported by refueling trucks.

1.5. RESEARCH OBJECTIVE

As discussed in previous sections, Shell is working to support the development of a global hydrogen market. The Shell team has already conducted an end-to-end value chain analysis that extends from the generation of electricity from solar and wind power to the production of hydrogen, ending with the supply of a growing number of customers in the heavy transport sectors. However, as discussed in paragraph 1.4, current roll-out plans still need backing by scientific research. Shell's plans to date have been based primarily on diesel data from their own stations. Current practice is to recalculate current diesel volumes by region to an ask for energy, enabling estimation of the energy equivalent demand for hydrogen. Modeling based on diesel refueling data reflects realistic behaviour as it is based on historical data. However, it also brings with it some disadvantages as (1) diesel refueling behaviour has specific characteristics and (2) Shell diesel data is not complete. The first is mainly caused by diesel prices varying across Europe and the second mainly rests on the fact that Shell is not omnipresent. The data is also limited to volumes per refueling, which can make it hard to pinpoint where heavy-duty trucks (in specific) are refueling. This introduces a set of drawbacks that need to be examined with a scientific approach:

1. Cross-border fueling due to varying excise taxes
2. Data gaps due to Shell's low presence in certain countries
3. Unclear what portion of fueling is done by heavy-duty trucks

The first drawback causes a bias in the overestimation of station size in countries with low excise taxes. It is common practice, especially for long haul transport, to refuel cross-border if fuel is cheaper there. This is clearly present in Shell's data, as it suggests enormous stations along the border of e.g. Poland. The second disadvantage causes general lack of knowledge of potential demand in certain regions. This leads to under- or overestimation of potential interest from Shell. The third disadvantage of using diesel data makes it hard to determine for Shell which locations are specifically interesting for heavy-duty trucks. As these type of vehicles will be the main target for hydrogen refueling, this is key in network planning.

The author of this thesis aims to investigate the optimal topology of a network of hydrogen refueling stations for heavy-duty trucks that takes these biases into account. Hence, this objective will be pursued using a method other than the translation of diesel volumes. The intended outcomes are recommendations and insights for station locations, along with their size. These stations should form a pan-European network capable of delivering sufficient hydrogen across the continent. To achieve this, not only the demand but also the supply of hydrogen must be considered. The cost of supply cannot be neglected, as a balance must be struck between the cost of supply and that of demand.

1.6. RESEARCH PROBLEM

The objective is to find the optimal topology of a European hydrogen refueling station network for heavy-duty trucks. The essence of the problem present is the balancing of costs that stem from demand and costs that stem from supply. On the one hand, Shell wants to minimize the costs for trucks to reach their refueling stations, and on the other hand, Shell wants to minimize the cost of supplying their stations with hydrogen. Hence, a station location problem like this asks for multi-objective optimization. As will be discussed in the literature study, studies either incorporate a two-step method or apply a tailored model that is able to handle multi-objective optimization. For this study, a two-step method seems in place: the refueling station locations can be modeled according to the demand from heavy-duty trucks whereafter they can be tested with respect to the relative costs of supply to their location.

2
LITERATURE STUDY

Before determining the research question and its sub-questions, a literature review is conducted. As a scientific approach is sought to deal with biases of company data, the literature surrounding refueling station locations is consulted. Much research has been done on optimally locating refuelling stations and there appear to be many methods for modeling such a network. Each model requires different inputs and produces different outputs. This section attempts to provide an overview of the most recent and relevant studies to uncover where research is lacking or what connections are missing, as well as to provide a first indication of methods for modeling. By grouping all the papers in Table 2.2 and indicating the scope and methods of the studies, the value of this paper as an addition to the scientific literature becomes clear.

2.1. LITERATURE REVIEW

To avoid bias of the author, a systematic literature review will be carried out. The research objective is reformulated to be a "topic", in which three clear concept groups can be distinguished. These concept groups are further specified into keywords, which make up the search strategy. The table below shows how the topic of this research is split up in these key words, and what search strategy is used for the literature study.

Table 2.1: In-depth explanation of search strategy

Topic	The **design of a network system** dealing with **uncertainty in demand** for **modeling heavy-duty trucks**		
Concept Groups	Design of a network system	Uncertainty in demand	Modeling heavy-duty trucks
Key Words	Network Design, Optimal Networks, Design Optimization	Demand uncertainty, Demand modeling, Demand forecasting	Heavy-duty trucks movements, Modeling heavy-duty trucks, HGV modeling, Demand modeling trucks
Truncation	(Concept 1) (Concept 2) AND (Concept 3) (Concept 2) OR (Concept 3) (Concept 1) AND (Concept 2) AND (Concept 3)		
Search Strategy	(Network Design Problem, Optimal Networks, Design Optimization) (Demand uncertainty, Demand modeling, Demand forecasting) AND (Heavy-duty trucks movements, Modeling heavy-duty trucks, HGV modeling, Demand modeling trucks) (Demand uncertainty, Demand modeling, Demand forecasting) OR (Heavy-duty trucks movements, Modeling heavy-duty trucks, HGV modeling, Demand modeling trucks) (Network Design, Optimal Networks, Design Optimization) AND (Demand uncertainty, Demand modeling, Demand forecasting) AND (Heavy-duty trucks movements, Modeling heavy-duty trucks, HGV modeling, Demand modeling trucks)		

All items found are assessed on their relevance and validity. Literature is only included if; (1) it is a peer-reviewed article; (2) the findings to include originate from the results, discussion or conclusion section; (3) the topic is related to network design, uncertainty in demand, and modeling heavy-duty truck movements.

2.2. CURRENT GAPS IN LITERATURE

This section will identify three gaps in the literature. Only literature that can account for increasing demand and specifically addresses hydrogen refueling will be reviewed.

2.2.1. LACK OF RESEARCH ON DEMAND MODELING FOR HEAVY-DUTY TRUCKS

As discussed in section 1.2, integration of the hydrogen fuel cell into long-haul trucks seems very likely. It seems inopportune, therefore, that refuelling station location research focuses primarily on vehicle passenger cars rather than on hydrogen demand from heavy-duty trucks [34–37]. Global emission standards for passenger cars have been in place for several years, a possible reason that the modeling of fueling infrastructure for these vehicles has been extensively studied. A methodological advantage of modeling demand from the point of view of the passenger car is that household origin-destination data can be used. This type of data is abundantly available and easy to process, since the vehicle always returns to its origin (the vehicle owner's home). This type of data can be processed in different ways to model demand for the different types of refueling location station models. For the set covering model, a location strategy can be coupled to a routing and scheduling strategy of household activity to locate refueling stations (e.g. Ref [35]). For the p-median model, the results can be combined with a geographic information system (GIS) based on one-way driving time from home or work to a station (e.g. Ref [34]). For flow-intercepting models, origin-destination travel demand data can be processed through a transportation simulator with the objective to place the refueling stations at locations which maximize the number of vehicles served, while staying within budget constraints [36]. Although passenger cars are the focus of these studies, the modeling approaches are still useful for heavy-duty trucks. A method that is also transferable, is that of Li *et al.* (2018) who aim to explore a mutual interaction between hydrogen fuel cell vehicle (HFCV) sales and the number of refueling stations. These studies show that there are several ways to find optimal station locations, based on data for passenger vehicles. However, the data for heavy trucks would not always be fully transferable. Modeling approaches for motorists are widespread, and to get an idea of the gap, it is necessary to know why they cannot be immediately converted.

Rose and Neumann (2020) suggest that "existing modelling approaches need to be adapted" for HDTs. They base this claim on the fact that the energy demand per individual refill will increases significantly for heavy-duty trucks and that it also brings regulatory and technical limitations for the capacity of the refueling stations. To this end, they did extend existing models and introduced an optimal development for refueling station locations regarding capacity restrictions. However, its use extends only to national use introducing a gap that will be discussed in the next section. The study of Nugroho *et al.* 2021 also models the location of the station with regard of heavy-duty trucks. These two studies already show that hydrogen location studies are increasingly emerging for heavy-duty trucks, and that the gap in this area is narrowing. However, the authors (who show a large overlap for both publications) identified shortcomings in the representativeness of the OD data for the HDT sector [3, 38, 39]. Kluschke *et al.* (2020) do include reliable origin-destination data of heavy-duty trucks in their study, as they use the number of individual HDT trips from a comprehensive road traffic survey in Germany. However, the scale of the study is only national. Research on modeling heavy-duty truck demand, involving more than one country, therefore jumps into a gap.

2.2.2. INATTENTION FOR EUROPEAN LEVEL OF AGGREGATION

The second gap that seems to emerge from the literature is that the European level of aggregation has not been taken into account or included in a case study. Many studies have focused on providing models that determine topology at the city level, without being scalable for continental use [41, 42]. Some efforts have been made to apply the previously mentioned models on a higher level [38, 40], but this number is very limited and was only at national level.

Section 1.3 discussed how the current implementation plans of a consortium, in which all relevant stakeholders are represented, are aimed at growth on a European scale. At a smaller scale, especially for heavy-duty trucks, hydrogen infrastructure is not cost-effective and unable to cover the international transportation market. Therefore, it is crucial that refueling station locations are also considered in a pan-European manner. It is indicated that the rollout plans will consume the European market iteratively, with a pan-European view starting in 2025. Thus, the addition of a European case study would add much value to the body of literature.

2.2.3. MISSING LINK OF CONNECTION TO HYDROGEN BACKBONE

The final gap discovered in the studies, is that the costs of supplying hydrogen by pipeline are hardly considered. As covered in section 1.4, supply of hydrogen to refueling stations can cost-effectively be covered by pipeline. However, the link to a hydrogen backbone to these stations seems to have hardly been present in the station location models in the literature. On the other end of the spectrum, Bae *et al.* (2020) proposed an optimization model for effectively deciding when and where to build HRS where mass supply of hydrogen is not possible. However most studies explore locations of hydrogen refueling stations that produce hydrogen locally [38, 41, 44, 45]. Methodologically, this asks for a very different approach. The local production of hydrogen affects its position due to the cost of supply in a very different way. The use of a flow-intercepting model is single objective, making it hard to incorporate multiple cost effects without adding a model or making thorough adaptations. A few account for multiobjective optimization within their model [41, 44, 45], others suggest a two-step model to deal with multiple objectives [46, 47]. However, none of these appear to have yet taken into account the cost of connecting to an existing hydrogen infrastructure network.

Nugroho *et al.* (2021) do compare fuel supply by pipeline versus on-site production. They compared the cost of hydrogen produced on-site with the cost of pipeline delivery from various centralized electrolyzer sites. However, their approach was such that the costs did not influence the choice of location, but revealed a difference in the cost of hydrogen production and distribution when location site was determined. Thus, the location decision does not take into account the costs of tapping into a hydrogen backbone. One study, a thesis of Hoekman (2021), does let the location of the hydrogen supply influence the choice of location. Both location decisions for refueling stations in terms of supply and vehicle routing decisions are considered. However, no implemented hydrogen pipelines are assumed to be present so the location decisions are based on road transport to two electrolyzer sites.

2.3. SUMMARY OF LITERATURE STUDY

The findings on the scope that the literature spans regarding this topic are summarized in table 2.2. The table provides an overview of where studies are lacking and where this paper will add to scientific literature.

Table 2.2: Summary of literature study

		Scope		
Reference	**Focus**	**EU**	**HDT**	**HBB**
Thiel	Pricing based location strategy			
Kuvvetli	Multi-objective and multi-period			
Kuby	The maximum volume of vehicle flows	√		
Kim&Kuby	Optimize locations with limited range and necessary deviations			
Lim	Heuristic algorithms for alternative fuel stations			
Miralinaghi	The effect of travelers' deviation to refuel on demand uncertainty			
Yildiz	Maximize the total flow covered	√		
Kang	Considers scheduling and routing decisions of individual vehicles			
Miralinaghi	Solution algorithms for multi-period refueling demand			
Rose	Interplay between HDT and HRFS that produce hydrogen locally		√	
Tafakkori	Designing sustainable refueling networks in urban areas			
Kuby & Lim	Extends flow-capturing models to optimal locations for refueling			
Shukla	Maximize the number of vehicles served, with budget constraints			
Hwang	Relax equal range and fuel level assumptions (add multi-class vehicles)		√	
Kluschter 2020	Incorporate adaptations in modeling refueling stations for HDTs		√	
Nugroho	Comparing different fuel supply scenarios (pipeline vs. on-site)		√	√
Hoekman	location decisions (for refueling stations) and vehicle routing decisions		√	√
Bae	Urban areas relatively far from a large hydrogen production site			
Tlili	Compares five hydrogen pathways (pipeline and truck options)			√
This paper	**European hydrogen refueling station network for heavy duty trucks**	√	√	√

In summary, the current literature lacks research that specifically addresses a heavy-duty truck refueling network on a European scale. In addition, there is a lack of integration of the hydrogen backbone into station location modeling. By incorporating the influence of a European hydrogen backbone on the location problem and by looking for HDT data for the demand modeling on a European case study, the gaps identified in all three sections are covered. The level of European aggregation is indicated by EU, heavy-duty trucks by HDT and the connection to a hydrogen backbone by HBB.

3
Thesis Setup

3.1. Research Questions

The research objective that emerges from the introduction is used as the basis for the research question. In addition, the question should attempt to fill the identified gaps in the literature. The following research question aims to do both:

Assuming supply from a European hydrogen backbone, what is the optimal configuration for a European hydrogen refueling station network for heavy-duty trucks based on expected demand?

To answer this main question, five sub-questions are formulated. These indicate the research activities and constructively set up the study:

1. ***What are relevant techno-economic parameters for a hydrogen refueling station network and its supply?***
   The purpose of this question is to find parameters relevant to the design and modeling of the hydrogen refueling network. It is expected that results will largely come from the technical system requirements for hydrogen supply and predictions for a future hydrogen markt in mobility.

2. ***What are the current plans and institutions for hydrogen refueling stations and how do they influence the network design?***
   The second sub-question serves to provide a broader context regarding hydrogen plans in Europe, with an emphasis on policy. The results will influence the system as derived from the first sub-question, as policy may dictate certain characteristics. As such, the results, together with the results of the first sub-question, will form a system description.

3. ***What is the expected European hydrogen demand from heavy-duty trucks over time, distributed over a highway network?***
   The third sub-question seeks to distribute demand of hydrogen over a highway network to enable placement of hydrogen refueling stations to cover this demand. This step should overcome aforementioned biases of current practice. The result should be suitable data on the presence of fuel cell heavy-duty trucks over time.

4. ***How are hydrogen refueling stations distributed over Europe in order to optimize coverage of European fuel cell heavy-duty truck demand?***
   The fourth sub-question addresses the first part of optimization. As a result, the distributed demand should be optimally covered by a network of refueling stations. This should reveal European hydrogen refueling potential and preferences and answer the fourth subquestion.

5. ***What is the influence of hydrogen backbone supply on the optimal locations of European hydrogen refueling stations?***
   The fifth question concerns the other side of the optimization problem, as the influence of supply per backbone will be sought. New station locations should emerge as a result of optimization of locations with supply in the mix. These locations should indicate the influence on supply for location determination over time.

6. ***What are critical uncertainties to the development of a European hydrogen refueling network?***
   The sixth and final question acts as a reflection, but also provides opportunity to include different scenarios. The uncertainties that remain as a result of assumptions and uncertainties about future developments will be discussed and tested through analysis. This concludes the study and allows clear recommendations to follow for Shell on what should be taken into account when determining the location of filling stations.

3.2. THESIS OUTLINE

This section outlines the structure of the thesis. The format of the thesis is discussed, as well as the coming chapters, each of which is briefly described. To straighten out the narrative, some chapters end with concluding remarks that clarify what has been done and where this thesis will go. This outline also provides guidance. Note that the numbering shows the chapter numbers and thus does not start at one.

4. **System Description**
   The purpose of this section is to examine what is currently happening in the sector and scope it down to what is relevant for this study. As such, a policy and supply chain analysis will be performed combined with research on fuel cell heavy-duty trucks and hydrogen pricing. The implications of the delineation will also be discussed to see what effects this scoping has on the study.

5. **Basis for Modeling**
   Prior to the theoretical framework, a basis for modeling is sought. First, performance measures for a network of gas stations are presented. These will be used to assess networks and methods to design them. For the modeling basis, theory on freight demand modeling is touched upon and a fitting methodology is chosen. The analysis of the regulations that influence the network also fall under this basis, as they are left out of the main model. It is separated from the theoretical framework to distinguish the model that is the core of this study and the preliminary work to enable this modeling.

6. **Theoretical Framework**
   The theoretical framework provides an overview of the methods within the theories of the core model: the station location model, network system design, and the two-step optimization. The section aims to delve deeper into these specific theories and their methods. Characteristics of the methods will be compared to the performance measures and criteria for the modelling itself. As a result, methods will be suggested as the right fit for this research.

7. **Freight Demand Modeling**
   The freight demand model will function as the basis for this study. The section will apply the methodology as put forward in the Basis for Modeling. The result will be a network representing the fuel cell heavy-duty trucks throughout Europe, over time.

8. **Regulatory Basis**
   The second part of the modeling basis is the regulatory basis. Regulatory constraints determined by the system description to be of possible influence on the network are dealt with. However, it is determined to not use these regulations as basis for the model but analyze them separately. This is done to exclude diesel data bias and ensure a free theoretical basis for the exercise. This is permissible as negotiations on this regulation are still ongoing.

9. **Station Location Modeling**
   The station location model represents the first step of the two-step optimization. In this chapter, the methodology will be presented followed by initial application of the model. The freight demand model will function as input, and the output will be a demand-optimal station topology. This topology will be used as input for the network system design.

10. **Network System Design**
    The second step in the optimization is the network system design. This chapter will apply the methodology proposed in the theoretical framework. The aim is to connect the demand-optimal stations to a source of supply in optimal fashion. This source of supply is determined to be an assumed hydrogen backbone. The resulting optimal network system design has demand-optimal locations, which are supplied in optimal fashion. This network will function as input in the two-step optimization.

11. **Two-step Optimization**
    The final step of the core model is the two-step optimization. This chapter applies the methodology for relocating the demand-optimal stations, with influence of supply. This influence should mimic a balance of costs of supply and demand. Note that this final step is called "two-step optimization" for the sake of reference, in reality the entire core model is the two-step optimization.

12. **Analysis of Results**
    After the two-step optimization, all modeling steps are complete. This leaves room for analysis of the models and their results. Since the modeling process was iterative, the analysis is performed in the same way. It is pointed out that not a single network design is the result, but the main findings collected during the modeling process are. The regulatory base is also analyzed separately to arrive at the key findings there. All key findings form the basis for the conclusions. This chapter also addresses key assumptions and limitations. The chapter ends with a validation of the key inputs. The latter feed the discussion of the final chapter.

13. **Scenarios and Uncertainty**
    Before drawing conclusions, this chapter elaborates on several scenarios. Uncertainties surrounding the model are discussed and specific areas of concern are highlighted. The most uncertain factors are then analyzed for different scenarios.

14. **Conclusions, Discussion and Recommendations**
    The final chapter presents the conclusions, discussion, and recommendations. The conclusions are drawn by repeating and answering the sub-questions one by one. The discussion makes some recommendations for proposing a follow-up study and provides feedback on the research and methodology. Finally, clear recommendations are formulated and presented to conclude the study.

To eliminate confusion about the structure of the thesis, figure 3.1 aims to support the storyline. The figure displays 8 chapters and shows how the basis for modeling is separated from the core model. Note that the system description provides input throughout the chapters.

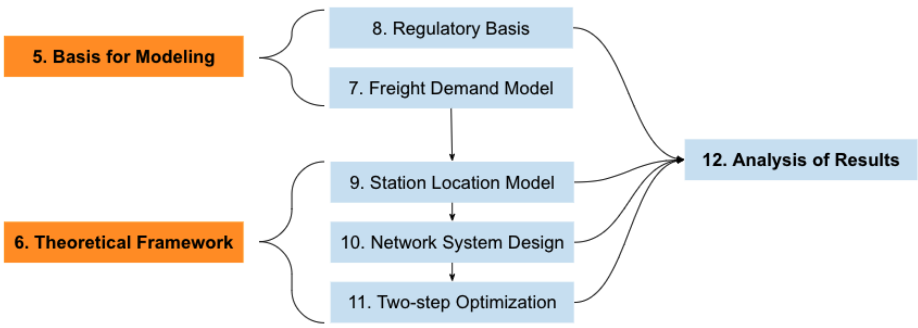


Figure 3.1: Supportive figure for thesis outline: the numbers indicate the chapter numbers

The basis for modeling supports the freight demand model and the regulatory basis. Note that the regulatory basis will not function as basis for this study, but will be analyzed separately. This is indicated by the line flowing directly to the analysis of results. The theoretical framework lays out the theories for the core model which is built up out of the station location model, the network system design and the two-step optimization. The output of these modeling steps will be analyzed separately in the analysis of results. Moreover, the station location model is fed by the freight demand model, as indicated in the figure.

4

System Description

The purpose of this section is to examine and reflect on what is currently happening in the sector in order to scope this research. There are many different technologies and methods for hydrogen refueling, and it is currently unclear which ones will gain the largest market shares. Modeling the system requires a description of what is there and, consequently, what part of it is taken into account. Relevant EU directives will be touched upon and technological trends, costs and limitations relevant to the study will be sought. As such, the focus of this research and the implications of the delineation will be described. This will lead to techno-economic parameters that must be considered in the modeling. To achieve that, this chapter is divided into five subsections. Three of which are supported by an appendix, to separate main issue from side issue.

First, this research's scope for fuel cell heavy-duty trucks will be described through a definition, an overview of characteristics, and a discussion of market penetration. Second, the supply chain as a whole will be analyzed. This will provide some parameters to model with, scope a distribution method and partially serve to provide an overview of what is out there. The third section will shortly touch upon predictions for hydrogen pricing "at the pump". Fourth, a full European policy analysis will be conducted to outline the EU's current plans and ensure that the scope identified is correct. Finally, these findings will be summarized in a table that presents the relevant parameters and findings for the rest of the research.

The research questions, along with their method, that the system description is looking to answer are identified in the table below:

Table 4.1: Research Goals of System Description

Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion	**Research Method**
What are relevant techno-economic parameters for a hydrogen refueling station network and its supply?	Literature review, policy mapping
What are the current plans and institutions for hydrogen refueling stations and how do they influence the network design?	Literature review, expert interview

4.1. Fuel Cell Heavy-Duty Trucks

As explained in the introduction, this chapter defines what is meant by a fuel cell heavy-duty truck (FC-HDT) in this thesis. Detailed research into the topic is presented in Appendix A, the paragraphs below present its key findings and delineation.

4.1.1. Definition and characteristics

The terminology on heavy-duty trucks (HDT) may vary in its exact wording and also in the resulting classification. Therefore, it is important to straighten out certain definitions and context prior to the rest of the system analysis. In this study, the EU vehicle category N3 and (semi) trailer category O4 are considered. As stated in Appendix A.1, heavy-duty trucks can be considered vehicles of 16 tonnes and above and only the EU falls within the scope. This entails 19.25m long trucks up to 42 tonnes with a maximum speed of 80 km/h as presented in Appendix A.2. The latter is based on the lowest maximum speed for these vehicle throughout the EU. The details in the Appendix consider a range of 830 km for 350 bar trucks with 50 kg of hydrogen aboard. This is based on a model of a potential FC-HDT. For 700 bar trucks with 85 kg of hydrogen aboard, a range of 1445 km is considered. Nugroho *et al.* present a CAD model of the potential FC-HDT which will be used as a basis for this thesis:

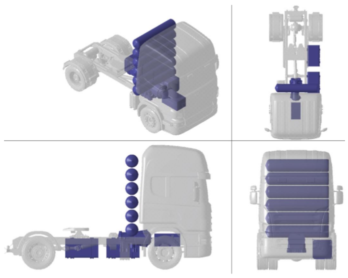


Figure 4.1: CAD model of potential FC-HDV [3]

Implications of delineation

FC-HDTs are the only customer of the refueling stations modeled in this study. They must be modeled separately to meet the study objective. In reality, hydrogen refueling stations may share their capacity with light-duty vehicles or other forms of mobility. These multimodal stations could cause a large drop in the levelized cost of hydrogen, as greater demand could be met with less investment. However, in order to provide specific insight into the truck market, they are not addressed in this study. For that same reason, all of heavy-duty traffic is housed under this one type of vehicle. In practice, more types of heavy vehicles, such as tractors, could refuel at the station. However, the focus of this study is on long-haul trucks for the research purpose previously indicated.

4.1.2. Predicted fuel cell market penetration

There is no single path to decarbonizing the trucking industry, and opinions vary on what the future will hold. However, the literature seems to agree on one viewpoint, which is that there is no silver bullet. Different types of propulsion will hold the future of transportation. To determine the infrastructure needed for a future fleet of FC-HDTs, it is important to factor in the potential penetration rate of fuel cells. In order to consider the different views and to clarify the position of this paper, a literature review on the subject is conducted and discussed in Appendix A.3. There appear to be few studies that focus exclusively on fuel-cell penetration rates for heavy-duty trucks and the European market in particular. One study, focused on e-fuels, includes a separation of freight and passenger vehicles [16]. E-fuels are gaseous and liquid fuels such as hydrogen, methane, synthetic petrol, and diesel fuels generated from renewable electricity. The majority of these e-fuels will be used for aviation, shipping and freight transport and they are currently in the phase of demonstration and very early market penetration. The study looks at different scenarios that describe the development of the share of powertrains and fuels for all transport modes in the EU. All their scenarios comply with the EU climate targets laid down in the 2030 climate energy framework and in the Energy Roadmap 2050 [COM/2011/885]. As such, each scenario analyzes the need for renewable energy to achieve an 80% or 95% reduction in greenhouse gas emissions in EU transport by 2050 (as opposed to 1990). However, each scenario dictates a different mix of "e-fuels". Their numbers, specifically for fuel cell heavy-duty trucks, are shown in the table below [16]:

Table 4.2: Market Penetration Scenarios for FC-HDTs [16]

Truck >12 t	2030	2040	2050
PtL-dom (%)	0	0	0
PtG-dom (%)	5	15	30
eDrives (%)	15	45	80

In the first scenario, It is assumed that internal combustion engines fueled with liquid fuels produced via power-to-liquid technology dominate all transportation modes. In the Power to gas dominated (PtG-dom) scenario, the focus is on power-to-gas fuels being increasingly used in electrified powertrains. It comprises PtCH4 for internal combustion engines and PtH2 for fuel cells. In the most rigid case, the eDrive-dominated scenario (eDrives), powertrains in all transport modes are electrified. To not base the rates for this thesis solely on one study, the results of Siegemund *et al.* (2017) are placed into a wider literature review in Appendix A.3. Kluschke *et al.* (2019) derived a summary of market diffusion studies of AFPs in HDVs, including their methods, main findings and policy recommendations. Siegemund *et al.*'s predictions for the market of alternative fuel powertrains (AFP) fall within 10% of the median of all studies combined and can thus be seen as valid numbers. While the eDrives scenario does not seem unfeasible today, a less fuel cell-positive scenario is being considered. Fuel-cell technology is developing quite fast and synthesis or methanation to liquid or gaseous options are in a very early stage. Nonetheless, for this study the moderate option towards fuel-cell is chosen. This is to remove potential bias of the author towards fuel cells and be consistent with Shell's expectations. Therefore, for this study, market penetrations for FC-HDTs of 5%, 15% and 30% are considered for modeling the next decades. It should be noted that the study does not take into account battery electric vehicles in freight transport because of their characteristics, but it is not the only study to do so [49–51].

IMPLICATIONS OF DELINEATION

The studies that emerged from the literature review are relatively "old". For most studies, scientific literature more than three years old is not such a problem. However, since this thesis deals with a rapidly changing market and political landscape, this is different. Nonetheless, the percentages chosen for consideration are based on scenarios that target a 95% emission reduction. This goal has barely changed and therefore the numbers are still useful. The greatest uncertainty that this study entails is the mix of fuels that will play a role in getting there. All technologies are still in the early stages and there is no real market for them yet, so the percentages are quite uncertain. As mentioned, the hydrogen and fuel cell markets stand a good chance. Nevertheless, the more conservative PtG figures are in line with Shell's expectations and therefore the best fit for this study. This implications should be kept in mind and possibly analyzed in a scenario.

4.2. SUPPLY CHAIN ANALYSIS

Now that the characteristics of the trucks to be modeled are documented, the refueling infrastructure is considered. To ensure that everything is taken into account, the entire supply chain is analyzed in Appendix B. At the end of each step in this analysis, a paragraph is dedicated to the "delineation". To keep focus on the main issue, only the delineation of parts relevant to this study are discussed in this section. It becomes clear what values are of importance for the rest of the research in terms of costs and techno-economic parameters.

To reason about which parts are of interest for this study, it is necessary to consider what factors are affected by the relocation of a station. In reality, each factor depends on location, as the costs of materials, labor, and energy vary by region. However, for this study, these costs are assumed to be equal. On top of that, a standard station and one form of supply is assumed. This oversimplifies the many different factors that can cause costs to fluctuate by (e.g.) station size, line of supply, and hydrogen production process. However, the complication of these differences would be too great for this study. Under these assumptions, the on-site supply chain costs remain the same regardless of the location of the station. The same is true for the production costs of hydrogen. Therefore, only distribution remains as relevant to the model's inputs. Figure 4.2 shows a schematic representation of the supply chain and HRS components, when being supplied by a pipeline. For additional information, a more detailed prescription of this part of the chain is included in Appendix B. In this main section, a brief description is given with reference to the corresponding section in the Appendix. Only hydrogen distribution will be highlighted in more detail, as its output is critical for modeling.

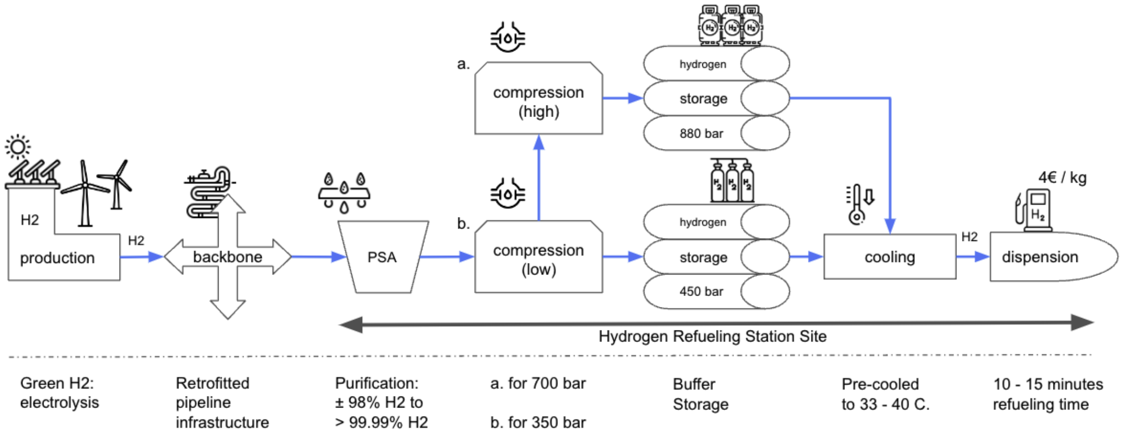


Figure 4.2: Schematic Representation of HRS components and supply

The hydrogen supply chain (as it is analyzed in the Appendix) exists of production (B.1) and distribution (B.3) before entering the HRS site. At the site, the hydrogen goes through a purification system (B.5) and thereafter, within the HRS, the supply chain consists of; a hydrogen storage system to meet daily demand (B.6); a high-pressure buffer storage system; a compressor that pressurizes hydrogen from the storage source pressure to the buffer storage pressure (B.4); a refrigeration system that pre-cools the hydrogen before it is dispensed into the HDT tank (B.7); and finally a dispenser that manages the flow of hydrogen to the tank (B.8). These are all accompanied by the necessary safety equipment and gauges.

4.2.1. Hydrogen distribution

The impact of distribution costs on station locations is clear: costs will be multiplied by the distance to a supply source. As previously identified, there is a gap in literature surrounding studies of hydrogen refueling station locations supplied by pipeline. Therefore, this is considered the only distribution method in the system of this thesis. As touched upon in Appendix B.3, this entails the inclusion of retrofitted gas infrastructure with their dimensions (transmission pipelines) and 98% pure hydrogen. Added to this are the addition of linepacking for storage (B.6), and the resulting need for purification of hydrogen (B.5). The hydrogen backbone is defined as consisting of transmission pipelines, compressor stations, control valves, and gas metering stations. The influence on the results for this thesis come from the techno economic parameters of the pipelines. Therefore, these are further visualized to make the scope clear and the remaining part of the section is devoted to explaining the hydrogen backbone and the piping system.

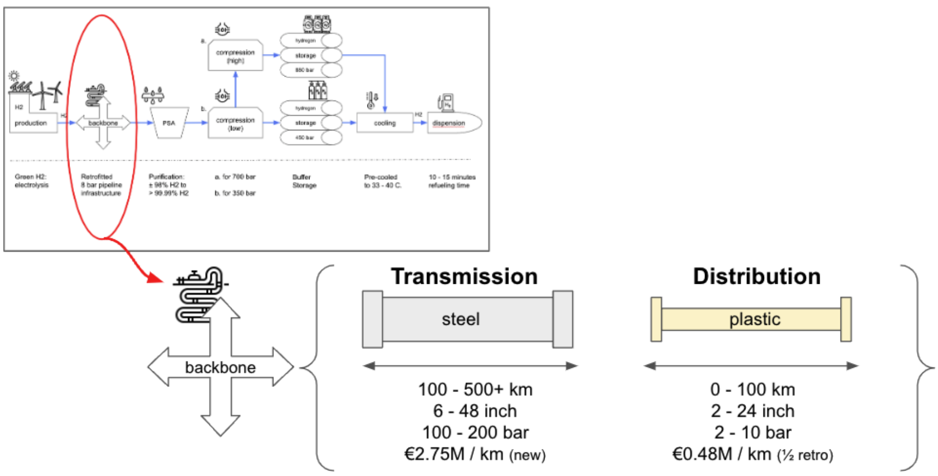


Figure 4.3: Schematic Representation of pipeline supply

HYDROGEN BACKBONE

Technical System

At the most fundamental level, pipeline delivery of gas happens through the gas flowing from higher to lower pressure. There are plans to develop a pan-European hydrogen infrastructure (a network of pipelines) which is called the European "hydrogen backbone" [52]. To get there, current gas pipelines can be retrofitted to enable transportation of hydrogen.

Hydrogen pipelines

The backbone will be made up out of transmission and distribution pipelines [53]. Hydrogen transmission occurs when hydrogen is transported from a central hydrogen production facility to a single point [54]. Hydrogen distribution occurs from a central hydrogen plant to a distributed network of refueling stations within a city or region. For this study, we can refer to transmission pipelines within the assumed backbone. The end connections from the backbone to the HRS are considered distribution pipelines. The hydrogen energy and capacity values in this section are based on the lower heating value (LHV) as it is customarily in energy system analyses and it enables for straightforward comparison between various fuels. For a more detailed and technical description of both types of pipelines, please refer to Appendix B.3.2.

Range of the backbone

For modeling purposes, it is necessary to know when an HRS is geographically "within the range" of a hydrogen backbone and with which capacity the pipeline should supply. These characteristics determine the length and diameter of the pipeline, which are the largest (non-location-specific) costs [3, 54]. Modeling supply by hydrogen pipeline has been done in several forms; Wulf *et al.* (2018) evaluated the environmental impacts of different hydrogen supply chains with differing demand and distances up to 400 km; Zhou *et al.* (2021) considered the transportation costs of hydrogen within 300 km; and Yang and Ogden (2007) research transportation costs where transmission distances varied from 25 to 500 km. It is difficult to find a specific value to determine whether an HRS is within "range" of a hydrogen backbone, as a network is modeled. The HRS are possibly interconnected, making a threshold of distance to the backbone itself invaluable. Therefore, the maximum distances from previous research (500 km) is taken to be "within the range" of the backbone.

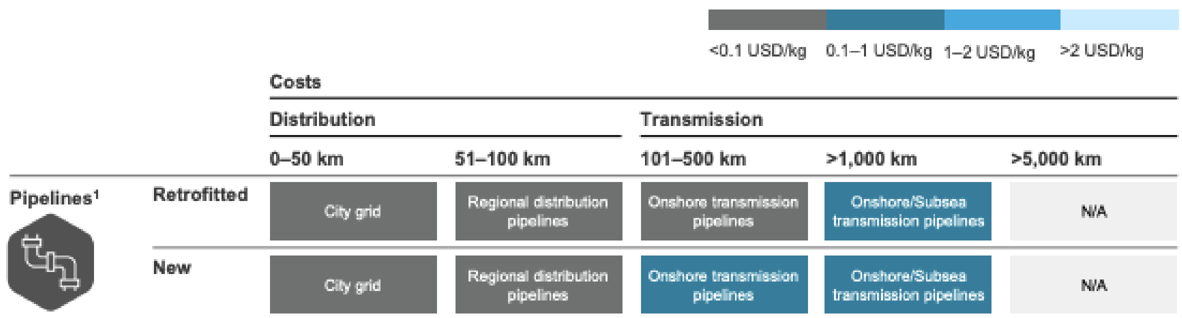


Figure 4.4: Pipeline distribution options: distribution prices assuming high utilization [4]

Figure 4.4 shows that the 500 km threshold coincides with the aim to keep the levelized cost of hydrogen delivery as low as possible. On top of that, the figure gives threshold values for when to assume pipelines of the distribution or transmission type.

Capacity of the backbone

As already stated, the capacity of the backbone will differ troughout the network along with its diameter. Boundary conditions for both transmission and distribution pipelines have been set. Elaborate calculations for capacity and corresponding costs are given by the technical brief of Adnan Khan *et al.* (2021), but transcend the scope of this study. For modeling purposes, it must be decided whether to consider a transmission or distribution pipeline. The backbone mainly consists of transmission pipelines and will carry the largest part of the transportation. However, this study assumes the presence of a hydrogen backbone and seeks to connect to it. It is assumed that the last 0 to 50 and 50 to 100 kilometers will be transported by distribution pipelines (figure 4.4). Hence, distribution pipelines are taken as the standard.

Analysis of costs

As for many large infrastructure networks, high investment costs account for the largest share of the total costs and inflict a barrier for creation. The total installed capital cost of a pipeline include not only materials for the pipeline, but installation costs, rights of way and miscellaneous costs [54].

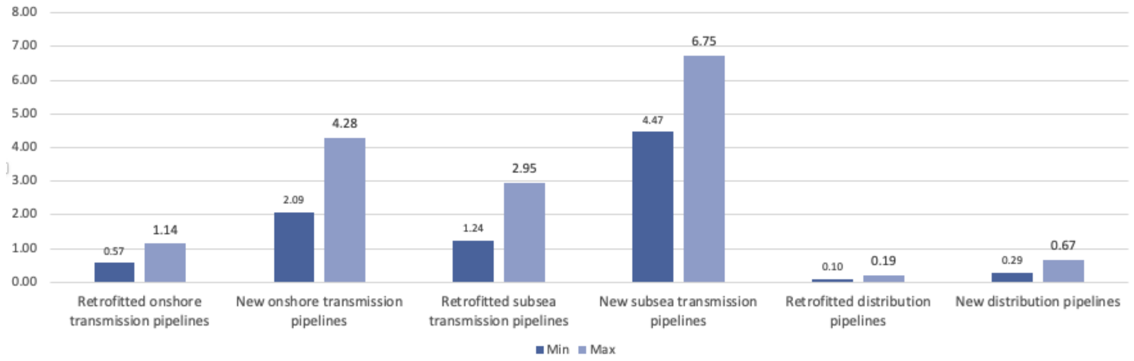


Figure 4.5: Overview of different pipeline costs in €M (based on [4])

Installation costs, rights of way and miscellaneous costs fluctuate greatly by region over Europe. However, for modeling purposes they are assumed to be equal. Adnan Khan *et al.* (2021) breaks down these costs and provides formulas to determine each. Their total costs estimates resemble the figures given in figure 4.5, hence it assumed all these elements are accounted for. The hydrogen backbone is defined as consisting of pipelines, compressor stations (if present), control valves, and gas metering stations. The figures also reside with the estimates of roughly €2.75 million for new pipelines of Wang *et al.* (2020). There is also validation for the cost of retrofitting, as Tezel and Hensgens 2021 estimate investment costs being reduced to €0.84 million per kilometer. Distribution pipelines are significantly cheaper than transmission pipelines given their smaller diameter and lower pressure.

The costs of retrofitting pipelines depend on a variety of factors including diameter and pressure, the quality of the materials used, the pipeline's overall condition, the existence of cracks, the social costs of construction, and other considerations [4]. Figure 4.6 breaks down the possibilities for the different type of pipelines:

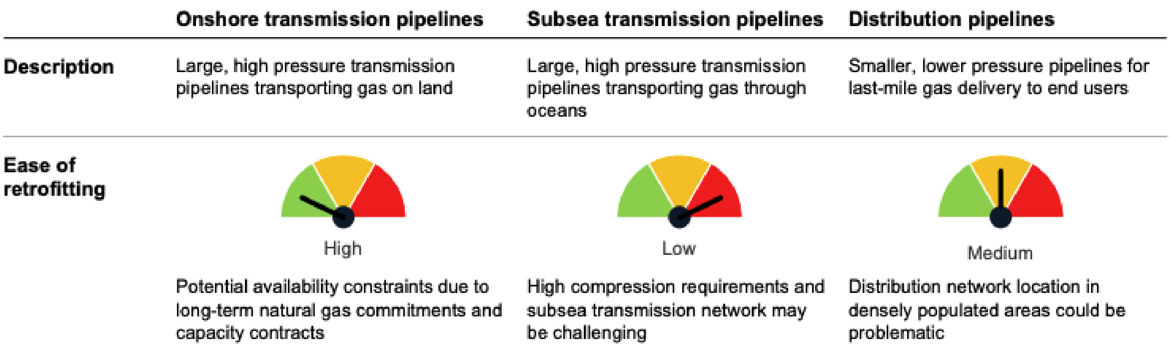


Figure 4.6: Pipeline distribution options: possibilities for retrofitting [4]

To calculate the cost of a hydrogen pipeline network, a cost figure per kilomter of pipeline is needed, as well as a figure for capacity expansion. Since distribution pipelines will be the subject of study, these figures are considered. The maximum values are taken as the starting point, as this excludes hydrogen-promoting bias. Since the retrofitting opportunities are ranked as medium (due to densely populated areas), the price per kilometer is taken as the average of the two numbers: € 0.48 M per km. The costs of capacity expansion are less straight forward. As previously discussed, complicated figures for costs of capacity are left out of scope. Therefore, a more simple exponential capacity factor that is used for natural gas infrastructure is considered: 0.6 [58, 59]. This beta represents a realistic value to indicate the cost advantages to building high capacity pipelines. It results in the effect that a connection with a capacity twice as much as another connection of the same length is not twice as expensive.

Levelised costs of hydrogen
The levelised costs of hydrogen transportation are estimated to be between €0.09-0.17 per kg of hydrogen per 1000 km, which indicates the method to be cost-effective over long distances across Europe [52]. This is determined using a conversion factor of 1 kg equalling 0.033 MWh (at lower heating value). However, this requires a certain level demand. Adnan Khan *et al.* (2021) established the rule of the thumb that "A demand of 1-1.2 tH2/day/km pipeline is needed to drive economic viability". When looking at cost optimisation for the total network, the concept of compression versus pipeline dimension and availability of the existing gas network are viewed as the main levers.

Relevant institutions
A pan-European hydrogen backbone will need to be backed up by politics in the form of regulations, incentives and subsidies. The implications will be large, and require holistic management: a review of the gas legislation, policy making on sustainable finance will all play a role in enabling this key European infrastructure. Currently, the European Commission has already announced integration of hydrogen infrastructure planning in their Hydrogen Strategy which will be discussed in section 4.4.2. Resemblence to gas legislation could have large consequences for the station location model. Similar regulation could cause that Transmission and distribution system operators (TSOs and DSOs) have the obligation to provide a connection to anyone who requests it. The costs of connection and transport would then be regulated and socialized in many EU countries. This could have the consequence that costs of connection to the backbone does not depend on the distance from the grid. This signals that the system could work very differently from road delivery in terms of costs. For modeling purposes, this is disregarded from this thesis.

Implication of delineation
It is clear that the given delineation of the supply chain is a major oversimplification of reality. In reality, the location of a station affects many factors and these should be considered as much as possible when determining a network topology. However, to model an overall balance between supply and demand, distance from a supply source must play a prominent role and this is most relevant for this study. Therefore, only the fluctuation of distribution costs is considered in the rest of this study. In this area, the fact that this study focuses exclusively on pipeline supply causes some discrepancies. By assuming a single source of supply, many opportunities for cost reduction in the supply line are overlooked. The costs of (e.g.) compression, storage or purification do vary by type of supply chain. Further optimization could take place when supply lines such as tube trailers or via liquid hydrogen tanks are considered. On the other hand, the fact that storage is left out leaves a large function of the backbone out of scope. A part of the use of a backbone is actually large-scale storage, so you can match demand with supply at all times. However, for the purpose of this study, these functions are not essential and are therefore not taken into account.

The focus for this thesis lies on the connection of the backbone to the HRS. Based on the delineation, this is assumed to be done by distribution piping. Distribution pipelines are scoped to cost relatively little, at €0.48 M per km. The costs of capacity expansion are determined by exponential capacity factor that is used for natural gas infrastructure: 0.6. It must be noted that these fundamental network design considerations do not fully mirror reality as costs will undoubtedly vary across Europe. Moreover, as for the market penetration, the costs of the backbone are uncertain and different studies point towards different numbers. Wang *et al.* 2020 suggest lower numbers than those outlined here in their study of the European hydrogen backbone. This is reinforced by discussion on which parts of the distribution will in reality be done by the type of pipe that is scoped. In the current gas infrastructure, plastic pipes go to homes at a pressure of 1 bar and an overpressure of up to 100 mbar. For this study the distribution pipes are assumed to cover a larger span of supply. The gas infrastructure is even set up in ring networks, for redundancy. It shows that resembling the gas infrastructure would lead to different aspects than are now scoped. Nevertheless, when looking at the system, the cost of distance to supply will increase with distance. The exact figures for distribution as scoped might not be fully accurate, but they fit well within the low and high expectations as shown in 4.5 and are thus assumed plausible. It is suggested that, as for market penetration, a scenario analysis is performed for possible validation.

4.3. HYDROGEN PRICING AT THE HRS

The cost of hydrogen is a topic often discussed in the literature. In contrast, the resulting selling prices are less discussed. Especially when looking for expected prices at the pump, little data is available. This section is intended to provide an understanding of where the price of hydrogen currently stands and what the expectations are for the future. The future price of hydrogen at the pump is important to the modeling as it determines the cost for an FC-HDT to drive to an HRS.

Current price

A study from the Hydrogen Council signals that hydrogen as a fuel will be cost competitive sooner than often expected [4]. In economic context, today's hydrogen market is in the kick-start phase (until 2025). From 2025 to 2035, a ramp-up phase is predicted, followed by a market growth phase until 2050 [60]. These first two phases are characterized by Hydrogen Valleys with regional and local hydrogen infrastructure. A hydrogen valley is "a geographical area where several hydrogen applications are combined into an integrated ecosystem that covers the entire value chain: production, storage distribution, and final use" [60]. Current hydrogen prices are researched by distributing a survey among these valleys. These numbers are displayed in figure 4.7. These mobility sales prices are currently far too high to be competitve.

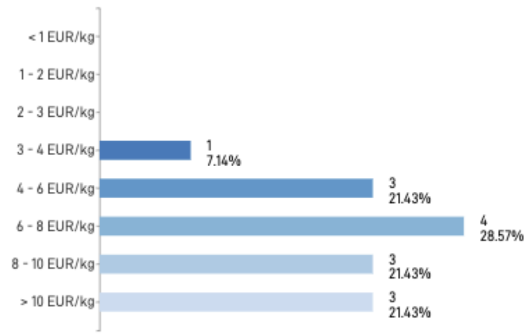


Figure 4.7: H2 mobility Sales price (share of Valleys) [5]

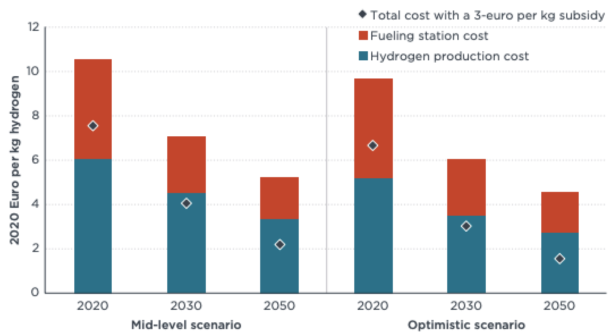


Figure 4.8: At-the-pump H2 price - avg. EU, on-site production [6]

Future price

Figure 4.8 shows the expected EU average at-the-pump price for onsite renewable hydrogen [6]. The prices are based on costs of hydrogen production (teal bar) and fueling cost (orange bar). On top of that, prices are indicated when a 3-euro per kg subsidy is provided (grey diamond). Their expectations for prices to lower in the future are based on technological improvements in both renewable electricity generation and in electrolysis. However, their main reason for assuming that these prices will decline are increased utilization rates. They assume utilization rates of the stations to be 30% in 2030 and 70% in 2050. Nonetheless, their estimate for hydrogen prices are significantly higher than the EU 2030 target of 1.8 euros per kg [61].

This shows that financial incentives are necessary to be viable. Subsidies and carbon taxes are often discussed. However, prices "at-the-pump" cannot be dependent on subsidies forever. When the market growth phase kicks in, in 2035, the hydrogen market should be transparent and liquid. From that moment onwards, price-setting should largely be governed by mechanisms of supply and demand. Therefore, this thesis assumes a subsidy, but one that will gradually decline from 2030 to 2050 as the market takes over. For the purposes of this thesis, the mid-level scenario was chosen. This suggests a hydrogen price of about 7 euros per kg in 2030, falling to about 5 euros in 2050. With a decreasing role of subsidies, a stable price for hydrogen can be set at 4 euros per kg for 2030, 2040, and 2050. The fully subsidized price for hydrogen in 2030 is assumed to be 4 euros, which can then be assumed to be the amount up to which the EC will subsidize.

IMPLICATION OF DELINEATION

A fixed priced for hydrogen over the next 3 decades seems implausible. However, the price of hydrogen must be such that it is cost competitive. At 4 euros per kg, the price of energy content of green hydrogen is comparable with gasoline [62]. It is assumed that green hydrogen is not a subject of environmental taxes, but only a subject of VAT. Due to low environmental impact of green hydrogen, this is a reasonable expectation.

4.4. Policy Analysis

In the analysis of the supply chain, and the fuel cell heavy-duty trucks, several regulations have already emerged. To make sure all relevant regulations are included in the rest of the report, this section aims to provide a regulatory overview from which a framework for this research can be derived. On top of that, in the first section EU strategies are discussed to get a broader view on the long-term goals of the union.

4.4.1. Strategies

Several strategies have been identified for alternative fuel infrastructure for HDVs. These strategies are further broken down and discussed in the sections below and Appendix C. First, the European Green Deal and regulatory incentives are discussed. Thereafter, the recent proposals of the REPowerEU Plan are taken into account. The Sustainable and Smart Mobility strategy are discussed in Appendix C, in conjunction with the Recharge and Refuel project. The NextGenerationEU plans are also analyzed in the Appendix. The Hydrogen Strategy is touched upon in the main text, as it is specifically relevant to this study. The national implemenation strategies are also shortly touched upon in Appendix C.

EU Green Deal

Compared to 1990 CO2 levels, the EU Green Deal aims to reduce greenhouse gas emissions by at least 55% in 2030 and reach climate neutrality by 2050 [63]. To get there, the EU is working on the revision of its climate, energy and transport-related legislation. They are doing so under the Fit for 55 package [64]. This package is a set of proposals for new iniatives and revisions of EU legislation to ensure that the policies are in line with the climate goals. These proposals include member states' emission reduction targets, an EU emissions trading system and a revision of the Renewable Energy Directive. The latter will be discussed in section 4.4.2 and proposes to increase the EU-wide renewable energy target to 40%, instead of 32%. Further relevant to this thesis are the proposals for alternative fuel infrastructure. The plans are included in the 2030 Climate Target Plan and the EC is working to push legislation on how to achieve these ambitious targets [65]. In June 2021, the EC made the EU Green Deal legally binding by adopting the European Climate Law [66]. The Alternative Fuels Infrastructure Directive (AFID) and the Transport-European Transport Network (TEN-T) Regulation fall within the framework of the Green Deal and will be discussed in section 4.4.2. The EC plans to support deployment of refuelling (but also public recharging) stations where gaps exist, with a focus on long-distance travel. This is beneficial for FC-HDTs, potentially enabling investments and subsidies in necessary infrastructure. However, a modal shift towards rail or inland waterways is also mentioned which may lead to an increase in competition of long-haul transport.

REPowerEU

In response to the global energy market disruption caused by Russia's invasion of Ukraine in 2022, the European Commission presented the REPowerEU Plan [24]. This plan entails a "massive scaling-up and speeding-up of renewable energy in power generation, industry, buildings and transport". This, with the goal to accelerate independence of the EU and give a boost to the green transition. As such, the EC proposes to increase the 2030 renewables target under the Fit for 55 package even further: from 40% to 45%. This ambition will create the framework for other initiatives. For the hydrogen market, the notable initiatives are: a target of 10 million tonnes of domestic renewable hydrogen production and 10 million tonnes of imports by 2030, increased sub-targets for specific sectors to accelerate the hydrogen market and Delegated Acts on the definition and production of renewable hydrogen. To achieve this, the REPowerEU communication included the Hydrogen Accelerator [67] which proposed €27 bn investment in renewable hydrogen by 2027 and beyond. This is aimed to be a direct investment in electrolysers and distribution of hydrogen in the EU. For the transport sector in specific, the Commission will present a Greening of Freight Package, aiming to significantly increase energy efficiency in the sector.

European Hydrogen Strategy

In order to implement the ambition of the European Green Deal [63] and the recovery plan [68], the European Hydrogen Strategy [69] sets out a vision of how the EU can transform clean hydrogen into a viable solution to decarbonize several sectors, over time. The goal is to install at least 6 GW of electrolytic renewable hydrogen electrolysers by 2024, followed by 40 GW of electrolytic renewable hydrogen plants by 2030. The strategy presents challenges and identifies the levers the EU can use. This results in a roadmap with actions for the years to come. The roadmap brings forward a concrete policy framework in which the European Clean Hydrogen Alliance (a collaboration between public authorities, industry and civil society) will develop an investment agenda. In addition to industrial applications, mobility is identified as one of the two leading markets for green hydrogen. The strategy prescribes first creating local hydrogen networks near production sites. Thereafter, transportation over longer distances is discussed influencing the TEN-E network. In this phase, the roll-out of a network of HRSs is envisioned. Therefore, review of regulation, as the AFID and TEN-T, are discussed in section 4.4.2. It is important to note, that CO2 emission standards regulation is seen as one of the largest drivers of the hydrogen economy as will be discussed in section 4.4.2.

4.4.2. Directives and Regulations

The overarching strategies and communications often refer to changes to specific directives and regulations. This section describes the most relevant of these. The directives and regulations are split up, based on their focus on infrastructure, vehicles or fuels.

Directives and regulations related to infrastructure

AFID

The directive that most directly supports the strategies discussed, is the Alternative Fuel Infrastructure Directive (AFID: Directive 2014/94/EU [70]). Introduced in 2014, the AFID provides a framework of measures for the deployment of alternative fuel infrastructure to minimize dependence on oil. The overarching goal is to reduce the environmental impact of transport, in line with the objectives of the Paris Agreement. By setting a minimum number of natural gas and hydrogen recharging and refueling points, the Directive sets requirements for infrastructure to cover and drive European demand. These requirements must be implemented through member states' non-binding national policy frameworks. To take the lead in a sector where there are many decisions to be made and standardization will lead to cost reductions, AFID also provides direction for standardization. For HDVs, AFID affects public infrastructure and it should be noted that private refueling points will also be important.

In the revision of REDII (as discussed later), AFID is set out to become AFIR. This implies that it moves from Directive to Regulation, and is thus directly applicable in the member states as it overrules national laws. It is ensured all alternative fuels under scope by detailed classification. The AFIR includes mandatory targets for the deployment of HRS along the TEN-T core and comprehensive network. Currently, it is assumed that by 2030, an HRS should be present (700 bar, 2t/day) every 150 km and in every urban node. Furthermore, liquid hydrogen should be available every 450 km.

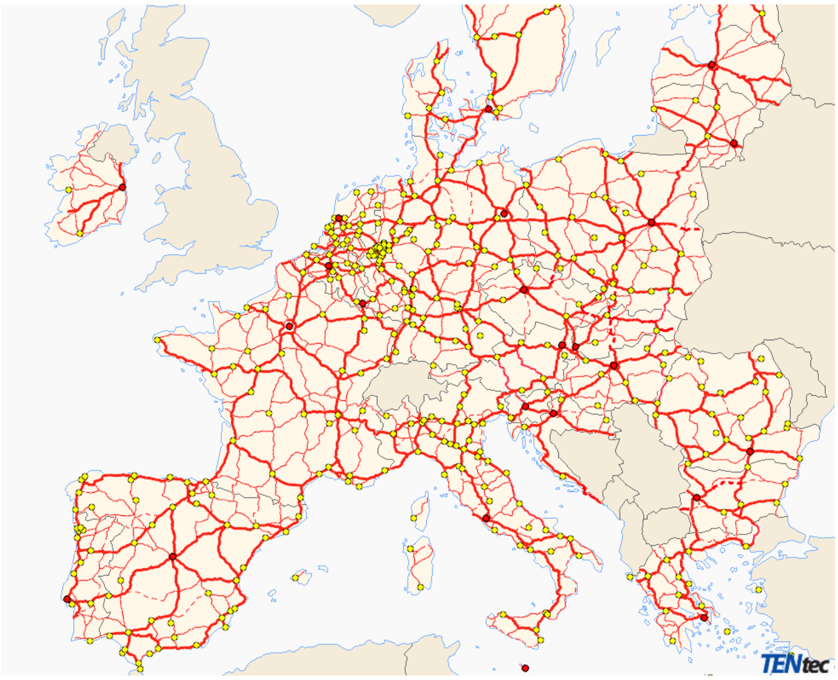


Figure 4.9: TEN-T visualisation: roads and urban nodes selected [7]

TEN-T Regulation
Very relevant to the research, is the Trans-European Transport Network (TEN-T) Regulation ((EU) No 1315/2013) [71]. The regulation establishes guidelines for the development of a trans-European transport network. Hence, it identifies projects of common interest and further specifies requirements, and sets out priorities for the management and development of the infrastructure. The wider aim is to develop a Europe-wide transport network by closing gaps and removing barriers. The network, as established in ANNEX I, is displayed on the figure below with support of the TENtec Interactive Map Viewer. The core network of ten corridors is to be completed by 2030 [7].

TEN-E Regulation
The Trans-European Networks for Energy (TEN-E) Regulation has the aim to link the energy infrastructures of member states. This framework for cross-border energy infrastructure planning was established to modernise and expand Europe's energy infrastructure [72]. It does so, with nine priority corridors for gas, oil and electricity and three priority thematic areas: smart-grid deployment, electricity highways and a cross-border carbon dioxide network. The TEN-E Regulation enables the EU to achieve core energy policy objectives by establishing rules for identifying Projects of Common Interest (PCIs) and ensuring timely implementation. It thus enables activities necessary for the energy transition and the achievement of climate goals. A recently approved proposal includes new and reused hydrogen transmission and storage infrastructure [73]. In addition, cross-border electrolysis facilities are discussed. The proposal to revise the TEN-E regulation aims to include hydrogen infrastructure and storage in the upcoming 10-year network development plans (TYNDP) to ensure timely implementation for the new market.

As there are many uncertainties (e.g. guarantee of origin of hydrogen) it is important to be wary of failing market structures. Therefore, at the moment, it seems best to establish the basic market rules for hydrogen based on the regulation of the natural gas market through the Gas Regulation ((EC) No 715/2009) and Gas Directive (2009/73/EC). These regulations impose non-discriminatory access to networks and unbundling rules that have created a successfull dynamic market for natural gas.

Directives and regulations related to Gas and Hydrogen
Renewable and low-carbon gases today face regulatory barriers for market and grid access. This represents a comparative disadvantage as opposed to natural gas and requires adaptation of the market framework. This lead to proposals for regulation and directives on the internal markets for renewable and natural gases and for hydrogen [74, 75] (COM/2021/803 and COM/2021/804).

Internal markets for renewable and natural gases and for hydrogen
The proposed initiative is strongly linked and complementary to the legislative proposals brought forward in the context of the Fit-for-55 package (e.g. RED and TEN-E). This initiative aims to enable a shift away from natural gas, and enable renewable and low-carbon gases to play the role necessary to achieve the EU's 2050 climate neutrality goal. The proposals include network planning, security of supply and storage, renewable gases in the existing gas infrastructure and customer protection. Currently, the regulatory framework does not address the deployment of hydrogen as an independent energy carrier through dedicated hydrogen networks. This means that there are no rules at EU level on tariff-based investments in networks, or on the ownership and operation of dedicated hydrogen networks [74]. Moreover, there is no single regulation of hydrogen quality. Consequently, barriers exist for the development of a cost-effective, cross-border hydrogen infrastructure and market. The regulations and directives proposals seek to address these deficiencies to support the uptake of hydrogen production and consumption. It includes a proposal for a system of terminology and certification of low carbon hydrogen and low carbon fuels.

Directives and regulations related to vehicles
Vehicle emission standards are one of the key drivers of the alternative fuel market in the European Union. For HDVs, standards for both GHG (Regulation 2019/1242) [76] and airpollutants (Regulation 595/2009) [77] are in force.

GHG Emission Standards
In 2018, the CO2 standard proposal for HDVs was put forward by the European Commission (EC) as part of Europe on the Move, a series of regulatory proposals [78]. In 2019, the regulation was formally adopted in the

Official Journal of the EU as Regulation EU 2019/1242 [76]. For HDVs, the CO2 standards are based on CO2 certificiation regulation (2017/2400) [79] and CO2 monitoring and reporting (2018/956) [80]. The regulation requires (for manufacturers) a 15% reduction for new HDVs for the period 2025 - 2029. For 2030 and beyond, the required reduction is set at 30%. In addition, an incentive mechanism has been established in the form of a system for awarding credits to manufacturers. Starting in 2025, this system will be replaced by a benchmarking system that will be established in the 2022 review. Relevant to this study, hydrogen will be included from that point on.

For N3 classified vehicles, the EC introduced a new segmentation based on max. weight, chassis configuration and axle type. Vehicle groups 4, 5, 9, and 10 correspond to the type of trucks are taken into account in this paper (indicated in table A.1). The vehicle groups are further divided into sub-groups to account for different use profiles, based on cabin type and engine power: (1) urban delivery (UD), (2) regional delivery (RD), (3) or long-haul (LH).

Air Pollutant Emission Standards

Regarding air pollutants, regulation 595/2009 sets out emissions standards for HDVs [77]. It covers exhaust emissions from all gasoline and diesel vehicles. More specifically, it covers "gaseous pollutants" (carbon monoxide, NOx and hydrocarbons) and "particulate pollutants" (components of the exhaust gas that are filtered out). The regulation establishes common technical requirements for the type approval of motor vehicles (and components thereof) with regard to emissions. In addition, the regulation includes requirements for in-service conformity of vehicles and engines, durability of pollution control devices, on-board diagnostic (OBD) systems, and measurement of fuel consumption and CO2 emissions. New standards are about to be implemented, as the Euro 7 standards initiative is expected to be approved by the EC in the first quarter of 2022 [81]. Future plans include the Eurovignette Directive, a toll system for trucks with a GVW of at least 12 tonnes.

Directives and regulations related to fuels

RED II

The most important directive in terms of fuels, is the Renewable Energy Directive II (RED II): 2018/2001 [82]. As the successor of RED, it sets a common framework for the promotion of energy from renewable sources. The directive targets for consumption of renewable energy for 2030 based on EU objectives. These targets are binding, and currently set on 32% renewable energy consumption. On top of that, it outlines requirements for member states' contributions to reach these objectives. Fuel suppliers are obliged to ensure that renewable energy takes up at least 14% in final energy consumption (3.5% being biofuels). How to include renewable fuels, like hydrogen, is still being determined by the EC and will be adopted through delegated acts.

Currently, RED II is already being gradually revised towards RED III under the Fit for 55 package. In this revision, the EU-wide binding renewables target is increased to 40% by 2030. Furthermore, the olbigated reduction target for fuel suppliers is seems to get changed to 13% for renewable fuels and electricy. Lastly, the subtargets for advanced biofuels (Annex 9A), are lowered to 2.2% in 2030.

FQD

The Fuel Quality Directive (FQD), 2009/30, sets targets for reducing the average greenhouse gas intensity of fuels [83]. It does not specifically address renewable energy deployment, but inherently has the same ambitions. The directive contains "technical specifications on health and environmental grounds for fuels to be used with positive ignition and compression-ignition engines, taking account of the technical requirements of those engines". Furthermore, the directive sets a target for the reduction of life cycle greenhouse gas emissions. Note that no decision has yet been made on extending the FQD target towards 2030.

4.5. RESULTING SCOPE AND FIGURES

Based on the description of the system, its delineation, and the general information presented, a set of system characteristics remains. This section aims to derive the information that will be useful in the modelling process of the thesis. Therefore, a table is presented summarizing the main delineations, followed by costs that will affect the modeling and finally possible political constraints, opportunities and problems. Chapters later in the thesis can use this section to refer to when using numbers or assumptions for modeling.

4.5.1. DELINEATION OF THE SYSTEM

The system description of FC-HDT characteristics, combined with the supply chain analysis resulted in a clear delineation for this thesis. Choices and assumptions have been clarified and substantiated. The results of this process are summarized in the table below.

Table 4.3: Delineation Table

Section	Delineation
FC-HDT	
Definitions	EU category N3 and O4
Characteristics	19.25m x 2.55m x 4.00m (l x w x h) Up to 42 tonnes 80 km/h Cylindrical type IV tanks for storage Range of 830 and 1445 km (350/700 bar)
Market penetration	5% in 2030 15% in 2040 30% in 2050
Supply Chain	
Production	Green hydrogen, 2.75 - 7€ per kg
Compression	20-30 bar ->30-40 bar (suction pressure)
Distribution	Hydrogen backbone: retrofitted gas infrastructure 98% pure hydrogen PSA for purification 48 inch diameter pipeline operating at 13 GW
Storage	Line packing and pressure buffering
Dispension	10 - 15 minutes refueling time

Furthermore, the system description prescribes a hydrogen price of 4 euros per kg, which will remain unchanged over the three decades. The latter is assumed to be the result of subsidies and market uptake. The price of 4 euros per kg should prevail to arrive at a diesel equivalent price of energy. It should be noted that the range of 830 and 1445 km is based on a fuel storage capacity of 50 and 85 kg of hydrogen. The market penetration rates are based on a scenario where the electric powertrain is not dominant but is present. It is assumed that the dominant share of fuels will be other renewable e-fuels. It remains to be seen whether this will be the future, but the figures are consistent with Shell's prediction and thus useful for this study. Further analysis of the supply chain is present in Appendix B. It should be kept in mind that production costs, compression, and storage via line packing are scoped to have no influence on the modeling.

4.5.2. COSTS FOR MODELING

The costs relevant for modeling are that of supply and demand. The supply costs are scoped to be those of connecting an HRS to a backbone by pipeline. This already has been scoped to be the CAPEX of semi-retrofitted distribution pipelines. The figures in table 4.4 reiterate how this lead to 0.48M€ per kilometer.

Table 4.4: Cost parameter table

Cost parameter	Unit	Low	High	This study
Distribution pipeline: CAPEX, new	M€ / km	0.29	0.67	0.48
Distribution pipeline: CAPEX, retrofit	M€ / km	0.10	0.19	
Depreciation period pipeline	Years	30	55	50

The costs to be considered for demand would be based on rerouting. The relocation of an HRS would result in trucks having to drive further or closer. Any kilometer these FC-HDTs would have to drive extra, would cost them extra fuel. Other aspects, such as the drivers' wages, are negligible and thus not considered. Based on subsidies and future market establishment, it was assumed that 1 kg of hydrogen will cost 4€ per kg in 2030, 2040 and 2050. Combined with the efficiency of a FC-HDT (17 km on 1 kg), this leads to the assumption, that it costs 4/17 € for every kilometer a FC-HDT has to drive extra.

4.5.3. POLITICAL CONSTRAINTS, OPPORTUNITIES, AND CONCERNS

In addition to delineation of characteristics and costs, the policy analysis serves to outline the context in regulatory terms. Strategies on European level lead to directives and regulations for different relevant sectors: infrastructure, vehicles and fuels. These international strategies may eventually lead to national strategies and together they form the political context. Important to the rest of the thesis are the constraints that define the rules of the game. In addition, there may be opportunities or problems in the game that models want to take into account.

Constraints:

- a hydrogen refueling station should be present (700 bar, 2t/day) every 150 km along TEN-T corridors
- a hydrogen refueling station should be present in every urban node

Opportunities:

- EU: 20 MTon of green hydrogen used by 2030
- obliged 32% renewable energy consumption in 2030 for EU member states (that is proposed to increase to 45%)
- 14% in final energy consumption of fuel suppliers
- national hydrogen strategies being rolled out

Concerns:

- potential for future failing market structures
- many uncertainties (e.g. guarantee of origin of hydrogen)

The constraints are both derived from AFIR and prescribe a basis for a network of hydrogen refueling stations. As such, they will be further considered as possible influences on the modeling process. The opportunities and concerns serve to provide the context and a picture of the future hydrogen market. The opportunities show the EU's commitment to renewable energy sources and the important role that hydrogen will play in this transition. Fit for 55 and REPowerEU show that ambitions are being raised at a rapid pace. The concerns focus primarily on the current market. Proposals for directives and regulation are presented but no proper regulatory framework is yet present to support uptake of a competitive hydrogen market.

4.6. CONCLUDING REMARKS

The system description is set up as a means of delineating the system within which this thesis will operate. In addition, it provides context on the market that will surround the network infrastructure. The future of the hydrogen market has many uncertainties, and for this thesis the delineation is based on scientific literature. In this fast-paced context and a still-unexisting market, literature quickly becomes obsolete and conflicting expectations are pervasive. For such reasons, many inputs can be debated and this should be kept in mind when interpreting the results. Based on this input, a basis for modeling will be set out in chapter 5. Thereafter, the results will be further used as input for modeling the optimal network topology in chapters 9, 10, and 11.

5

BASIS FOR MODELING

This chapter provides a basis for modeling. This basis is separated from the theoretical framework to clearly distinguish between the model that is the core of this study and the preliminary work to enable this modeling.

First, performance measures for a network of refueling stations will be established. This will allow the selection of a methodology that enables optimization in terms of these measures. Optimization is not possible in the absence of clear criteria. At a later stage, the networks resulting from the model can also be analyzed using these measures. Second, freight demand modeling is discussed. A freight demand model is intended to remove the bias of Shell diesel data and should function as the aforementioned "truck traffic counts". The output should allow for further analysis of the network. Theory within freight demand modeling is outlined and different methods are discussed. In accordance with the purpose of this study and the availability of data, an appropriate method will be chosen. The final section of this chapter will discuss implementation of the regulations that have been identified as affecting the network. By discussing these three steps, a clear basis for modeling the optimal network will emerge.

5.1. PERFORMANCE MEASURES FOR A NETWORK OF HRS

As indicated, performance measures must be established to get an objective definition of what makes a network good. These criteria will be used to set up proper methodology and later on analyse the results. To find them, general criteria for network planning are sought and objectives of other studies are consulted.

For general network planning, a proper design must; provide reliable and secure supply; meet safety and environmental standards; and optimise equipment utilisation and network losses [84, 85]. These criteria are not all equally relevant to this study, but they do provide a basis. In station location determination, securing supply and optimising utilisation are most relevant. In this light, one study, specific on refueling station networks, aims to build as few stations as possible while still adequately serving consumers [86]. This is a first indicator of a search for a balance. They point towards the high cost of constructing alternative fuel stations. Therefore, they conclude that it is essential to coordinate the locations of stations to allow maximum consumer coverage. Furthermore, the utilization rate is specifically mentioned as undeniably important to the financial viability of a station, especially in the initial phase. Another study alligns on this goal, as it minimizes the capital cost for constructing the refueling infrastructure [87]. They oppose this aim with the goal to maximize the total vehicle-miles traveled (VMT) per year covered by the stations. In the same line, Honma and Kuby (2019) aim to minimize the percentage of infeasible demands (demand points not covered by the network of stations). Furthermore, they report the weighted average travel time to a refueling station as a performance measure. The objectives of these studies can be summarized and translated into performance measures (in bold) as follows:

1. To minimize **CAPEX** (number of stations and their supply)
2. To maximize the **utilization rate** of the HRS
3. To maximize the **total vehicle kilometers covered**
4. To minimize the **weighted average distance** (travel time) from the located demand to any HRS

It should be noted that different preferences appear to exist for each stakeholder. Bai and Zhang (2020) break down preferences into those for governments and those for private partners. Governments seek to ensure the quality and coverage of the infrastructure, while private partners are concerned with the financial risks. For an optimal network for all parties, a balance between the two must be sought.

The four performance measures are already indicative of the search for that balance. The balance can be simplified by a general search for balance between costs for supply and demand. Supply is represented by operator costs (CAPEX and utilization rates) and demand is represented by costs of coverage (weighted distances and vehicle kilometers covered). The theories, and subsequent methodology, created for this study should set out to optimize this balance.

5.2. FREIGHT DEMAND MODELING

Prior to any optimization, the basis of the network must be established. This base must be constructed from data that relates to freight movements (as opposed to diesel refueling data) to achieve the research goal. Therefore, a method must be found within the theory of freight demand modeling.

In recent decades, various modeling methods have been developed and applied to predict freight movements between and within economic areas. The main (dis)advantages are found in data demand and aggregation level. A synthesis of approaches is given by the literature [90]:

Trend and Time Series Analysis
This type of analysis involves the longitudinal extrapolation of historical trends in freight activity. They range from simple growth factor models to complex models fit to time series data. The advantage is that these analyses are easy to perform and do not require a lot of data. The general disadvantage is that they are usually aggregate in nature and do not include explanatory variables.

Elasticity Methods
Elasticity methods measure the sensitivity of demand to a specific variable. Measurements of elasticity vary depending on the available data and application context. These type of methods are useful for general order-of-magnitude estimates of transportation. However, the inability of this method to simultaneously take multiple factors into account make it difficult to apply for this research.

Network Models of Economics and Logistics
Network models focus on the behaviour of shippers and carriers. In this way, they attempt to capture both supply and demand that generate freight movement. The models have great potential for intercity freight flows, but for modeling more aggregate levels, the methods have data requirements that are too intensive.

Disaggregate Models
The disaggregated models can be seen as those that focus on mode choice behavior, since it has its origins in microeconomics. These models usually take the form of logit models for utility maximization. A disaggregate approach that is specifically atuned to the view of a firm is the "inventory approach" [91]. Variables related to production are treated as endogenous decisions along with the mode choice. As one can imagine, the difficulty in using disaggregate models is obtaining suitable data.

Aggregate Demand Models
The essence of aggregate demand models is that they are based on estimated traffic flows rather than the number of individual trips. For this reason, their output is more aggregate than disaggregate models. There are several classes worth mentioning [92]; (1) the total flow approach, which establishes the relationship between industry output and demand (utilizes aggregate time-series data); (2) development of a modal split, establishing the share of trucks in traffic (its advantage being taking more than a single aspect into account); (3) the generation of synthetic origin-destination (OD) matrices from truck traffic counts. The latter has the advantage of working well without detailed data.

Economic Input-Output Methods
Economic input-output methods use economic input and output indicators. These determine the level of economic activity between regions, which in turn drives the demand for freight transport. Inputs are basic resources that add value through economic activity, the outputs are the quantity of goods demanded by type, location, and temporal frame. The outputs can then be converted to get estimates of the freight transportation demand.

5.2.1. FIT FOR THIS THESIS

As established in the research objective, modeling freight demand requires finding a method that is capable of modeling heavy-duty trucks to get to local hydrogen demand. It should be known where trucks are at a moment in time. In this sense, trucks could be considered as demand points or flows. On top of that, this demand should be distributed over a network of important European highways (corridors). Therefore, the method must be able to deal with the European level of aggregation.

Due to the necessary level of aggregation, network models of economics and logistics and disaggregate models can immediately be deemed impracticable to meet the requirements of this thesis. As already indicated, elasticity methods are also no good fit for the research due to their focus on the effect of one specific variable. This leaves trend and time series analysis, aggregate demand models and economic input-out methods.

These three models seem fit to handle the aggregate character of the demand. Out of these approaches, aggregate demand models seem to have the best fit. Generation of synthetic OD matrices would yield the best results as they would be based on actual numbers of trucks on the road. However, this data might be hard to come by on such a high level. The same can be expected for trend and time series analysis, as they require a previous study on truck traffic flows over the European highway network to extrapolate. This is where economic input-output methods would have an advantage, as traffic flows would be created synthetically, based on economic activity. In general, the European Union has well-organized data on incoming and outgoing goods. The demand would be less accurately placed since the data is synthetic, but the data search is easier.

DATA SEARCH

This section searches for data that fits (one of) the methods, and a choice for a method is made accordingly. To get there, first the data requirements to meet the objectives of this thesis are specified:

- data must be abundant enough to allow for modeling at the European aggregation level.
- data must be specific enough to distinguish different regions within Europe (sufficiently disaggregated to take into account a subnational level).
- data must be able to be recalculated to account for the increasing demand for hydrogen by HDTs over time.

To satisfy these requirements, the three models would require different inputs:

1. The envisioned aggregate demand models need pan-European truck traffic counts as input.
2. The economic input-output methods would need incoming and outgoing goods per European region (on a sub-national scale).
3. The trend and time series analysis would need previously worked out models.

MODELING DECISION

To the author's knowledge, there are no data sources that cover the full European scale for truck traffic counts. Moreover, no previously developed models have been found that allow for trend and time series analyses. However, counts of incoming and outgoing goods per European region are present [93]. The source provides data on the annual road freight traffic classified up to NUTS 3 [9] for loading and unloading. The NUTS (Nomenclature of territorial units for statistics) classification is a hierarchical system that divides up economic territory of the EU. The current NUTS 2021 classification is valid from 1 January 2021 and lists 92 regions at NUTS 1, 242 regions at NUTS 2 and 1166 regions at NUTS 3 level. Therefore, NUTS 3 regions are appropriate for distinguishing subnational regions at a scale small enough to generate synthetic data. Hence, **economic input-output methods** were chosen as the fit for this thesis.

This choice entails certain consequences, as the number of trucks on the road will depend on the number of goods to be traded. Within these numbers, changes caused by the energy transition will play their part as well. The type of goods transported by road was collected according to the 24 groups of goods following the "Standard goods classification for transport statistics 2007, NST 2007" [93]. This classification includes (e.g.) products of agriculture, mail, parcels, metal ores, but also crude coal, petroleum and natural gas. As a result of EU regulations, products such as the latter will see a sharp decline, while other goods may see an increase. As relevant as these effects are, they are not considered in this thesis because of their complexity.

5.3. REGULATORY BASIS

Regulation that needs to be discussed prior to creating a model is previously identified in the system description, section 4.5. Any relevant constraints, specific costs or other system characteristics that emerged must be taken into account or analyzed. The policy analysis outlines two constraints relevant to station location determination:

- an HRS should be present (700 bar, 2t/day) every 150 km along TEN-T corridors
- an HRS should be present in every urban node

These restrictions stem from the AFIR, which is currently being negotiated. Once implemented, the AFIR will be binding since it is a regulation. If negotiations are finalized, the basis of any HRS network should be the constraints inflicted by the AFIR. However, as the regulation is still under negotiation, analysis of the proposed station sites is more useful. At this point, implementing the sites as the basis of the network itself would lead to false assumptions, as negotiations are still in progress. Instead, spearate analysis could lead to policy recommendations that can inform the negotiations.

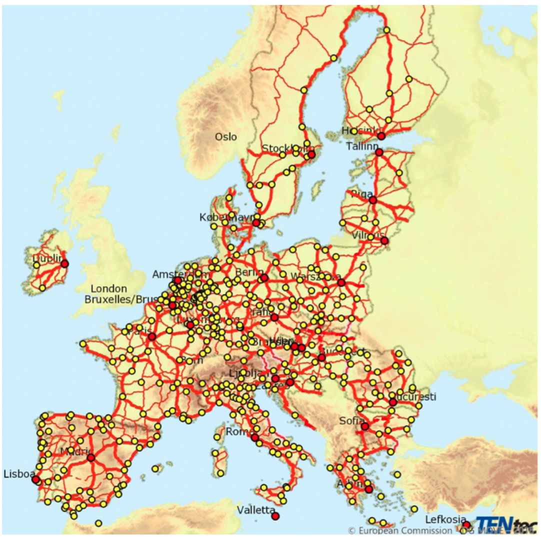


Figure 5.1: Urban nodes as defined by AFIR

A side effect of the AFIR as the basis of the model would be that stations would not be optimally located. The station locations that would result from the model would be influenced by policy, changing the insights from the study. Excluding the AFIR constraints provides a free theoretical basis.

5.4. CONCLUDING REMARKS

This section aimed at establishing a basis for modeling an optimal network topology. In that light, four performance measures are drawn up. These performance measures will guide the theoretical framework in section 6. Furthermore, they will be called upon in the analysis of the results (12). Second, a selection of a freight demand model that can meet the research objective of this thesis was made. This methodology will be applied in section 7. Finally, the possible regulatory basis for a network of refueling stations is discussed. It is decided to perform an analysis on the network as prescribed by the AFIR but not use them as the basis for the modeling. This is based on the fact that the regulation is still under negotiation and constraints would impede freedom of the model. The methodology for this analysis will be discussed in section 8.

6

THEORETICAL FRAMEWORK

In this chapter, a theoretical framework is outlined. The framework is intended to provide a skeletal structure within which appropriate methods are proposed. To arrive at theories that fit the objective of this study, the performance measures of the modeling basis are needed. These measures point towards the optimal network of refueling stations and give direction for the theories. Ultimately, the methodology should allow for the optimization of performance measures in the form of a balance. Theories are chosen accordingly in section 6.1. In section 6.2 and 6.3 pros and cons of methods within these theories are shortly discussed. Each section ends with a description of the respective fit for this thesis and a judgment as to which method is the most appropriate. These discussions also take into account the availability of data.

6.1. TWO-STEP OPTIMIZATION

The modeling basis prescribed a balance between costs for supply and demand. It stated that the four performance measures are already indicative of the search for that balance. Supply is represented by operator costs (CAPEX and utilization rates) and demand is represented by costs of coverage (weighted distances and vehicle kilometers covered). The theories, and subsequent methodology, created for this study should set out to optimize this balance. As a reminder, the four performance measures (pm) are restated:

1. To minimize **CAPEX** (number of stations and their supply)
2. To maximize the **utilization rate** of the HRS
3. To maximize the **total vehicle kilometers covered**
4. To minimize the **weighted average distance** (travel time) from the located demand to any HRS

To find the balance between supply and demand, a two-step optimization is proposed. The use of such optimization allows for the search for balance, as one can be optimized followed by the other. Many other studies on the same topic effectively rely on the same type of optimization [3, 38, 46, 55]. The research objective of this thesis initially focuses on the demand side, as truck traffic counts were to form the basis.

Station location modeling
Therefore, the first step will be to create a station location model that determines the locations demand-optimally. Existing Shell diesel stations are excluded from the model to ensure a free theoretical basis for the exercise. The focus should be on the coverage of the network and the average weighted distance from the located demand to each station (pm 3 and 4). However by aiming at an optimal number of stations to be built, utilization and CAPEX of stations can also be targeted (pm 1 and 2).

Network system design
For the second step, the station locations will be connected to a supply source (the backbone). This connection to a source can itself be implemented in an optimal manner, resulting in an optimal network system design. By calculating the cost of this network (CAPEX of piping) locations can be further optimized in terms of supply (pm 1).

Two-step optimization

The final step is to optimize the topology by balancing the costs of supply and demand. This will be done by balancing a cost function. Note that this final step is called "two-step optimization" for the sake of reference, in reality the entire process is the two-step optimization. The structure of the complete two-step optimization is visualized in figure 6.1.

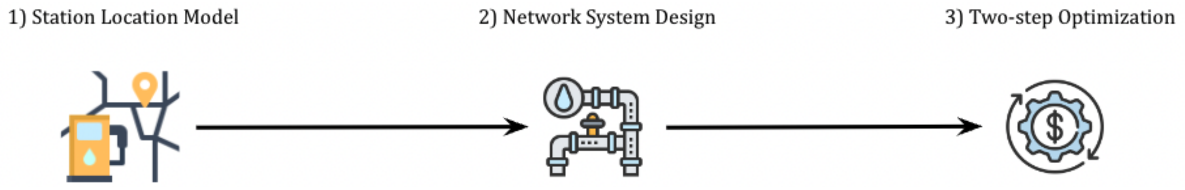


Figure 6.1: Skeletal structure theoretical framework

Balance

To find the right methodology within these overarching theories, each step must affect the performance measures to achieve the intended balance. For this reason, the performance measures can be tracked during each step of optimization. Starting with demand, moving to supply, finally ending in balance. The process of balancing, measured by the performance measures, is shown in figure 6.2.

Figure 6.2: Search for balance (SLM = station location model, NSD = network system design, 2SO = two-step optimization)

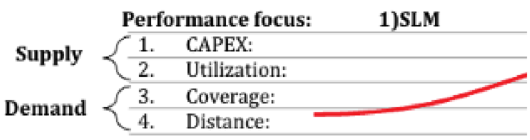
Pros and cons of methods within these theories are shortly discussed followed by a verdict on the respective fit for this thesis. Data availability is also taken into account in these discussions. Note that the output of the previously discussed freight demand model (5.2) will function as input for the optimization. First, theories on station location modeling are explored in section 6.2. Second, in section 6.3, theories on Network System Design are discussed. Finally, the two-step optimization is discussed in section 6.4. The first two theories are explored by a short literature review and unfit methods will be disregarded. The two-step optimization will be based on balancing a cost function as a result of the first two steps. No literature review is required, but a brief discussion of the theory will be given. This will be applied in chapter 11. Eventually, the methods that fit the purpose of this study best will be elaborated and applied in chapters 9, and 10.

6.2. STATION LOCATION MODELING

As indicated, the station location model will function as a first step in a two-step optimzation. The results sought are demand-optimal locations, in terms of the four performance measures. To find the right method within the theory of station location modeling, the performance measures are scored to fit the balance as displayed in figure 6.2. This resulted in the following table and visualisation:

Performance measure	Focus SLM
1. CAPEX	+ (# stations)
2. Utilization	+
3. Coverage	+++
4. Distance	+++

(a) Performance measures Station Location Modeling



(b) Performance focus Station Location Modeling

Figure 6.3: Performance focus Station Location Modeling

Figure 6.3b reveals what is asked of the methodology which is translated to figure 6.3a. It should be able to optimize the coverage of the network and the distance of the trucks to the refueling stations first. Thereafter, further optimization in terms of supply is welcome.

The main body of literature dealing with research on refuelling station topology does so in the form of a "station location problem". The station location problem can be considered as a network design problem when wanting to determine the supply of relevant facilities. The relevant facilities would in this case be represented by refueling stations. Lin *et al.* (2020) distinguish five different type of models that are used in the literature to tackle such station location problems: (1) the set covering model ; (2) the maximal covering location model (MC); (3) the p-median model; (4) the p-center model ; (5) the flow-intercepting model. A short explanation of each model is provided in Appendix D.

6.2.1. FIT FOR THIS THESIS

The choice for the best fitting station location model for this thesis is complex but should be in line with the strategy of balancing demand and supply. It is previously determined that the initial step will lean towards demand. Of the five types of models used in literature, the first two (set and maximal covering) models do not consider demand and are therefore not suitable for the goals of this thesis. With the same line of thought, the p-center model is deemed not a good fit with the goals of this thesis. The model is not directly linked to demand calculation and thus not suitable for determination of refueling station locations. This leaves p-median and flow-intercepting models.

The choice of the model is further influenced by the input that stems from the freight demand model. The p-median would require trucks to be represented by demand points, where the flow-intercepting model would handle a truck flow over a network. Flow intercepting models seem to be the best way to deal with this output, as economic input-output methods predict flows between regions. However, on a European level, the computational times may become too expensive to find an optimal solution. The more simple p-median model would not encounter such problems but there is one major problem inherent in using the p-median model: as it is a discrete location model, the set of possible station locations is a given finite set. This goes against the objectives of this study, as the intention is for the station location to be determined by demand and not predetermined.

Therefore, additions to the list of Lin *et al.* (2020) are suggested for this thesis: a continuous facility location model (ConFL) and k-means models. ConFL is the variant of the p-median model in which a facility may be opened at any point of the Euclidean space and has applications in clustering. The clustering effort makes these models closely related to the more widely discussed k-means problems [95]. The difference with k-clustering is that ConFL does not impose a predefined number k of clusters. Instead, it fixes a penalty for each created cluster. To make the choice between these two models or a flow intercept (FI) model less arbitrary, as all are able to take into account demand, a deeper analysis is required.

FURTHER ANALYSIS OF FI, CONFL AND K-MEANS

Flow Intercept Models (FRLM), k-means models, and continuous facility location models (ConFL) are considered appropriate for this thesis and will be further discussed to check which model fits this thesis best.

The flow-intercepting model.

Flow-intercepting models are focussed on finding the number of refueling stations that maximize the refueled traffic flow between origin-destination (OD) pairs, considering a limited driving range of vehicles. The demand is associated with the traffic flow, which can be calculated in a dynamic way, making it very good for application to refuelling stations. Within this field, the flow-refueling location model (FRLM) [96] is a popular method as it does consider tanking more than once. It is said to be the most dominate option in road transportation research [97–99]. The FRLM is path-based, and locates p stations to maximize the number of vehicles on their shortest paths that can be refueled. The main objective of the FRLM is to minimize the number of facilities needed to cover a given demand share, or to maximize the vehicle trips covered when locating a predefined number of stations in a network. One model has been found in literature that is focused on hydrogen refueling station locations: the NC-FRLM [38]. It determines the minimum number of capacitated refuelling stations to serve a pre-defined share of the vehicle flow.

ConFLP

General facility location problems aim to determine sites for new facilities with respect to a set of fixed points. In this case these would be the refuelling stations as opposed to the demand points. The Continuous Facilition Location Problem is a natural extension of the Facility Location Problem [95]. Location problems become continuous when the underlying space, for refueling station as well as the demand points, is a continuous one [100]. ConFLP starts with opening stations one by one, each at a certain cost, connecting each datapoint to a station. The objective is to minimize the costs of opening stations and also that of connecting datapoints. Therefore, these models were appropriately referred to as "site generation" models. No knowledge of candidate sites is assumed, since the model is supposed to generate the appropriate sites for the stations.

K-**means**
K-means algorithms aim at clustering [101]. Clustering is an unsupervised machine learning method, that divides items into groups, such that objects in a group are more similar than those outside the group. This similarity is often defined by distance. In the case of this thesis, that is the right fit, as datapoints should be clustered towards stations based on their proximity. The distance (called Euclidean distance) between the station and the datapoint is calculated by using Pythagorean theorem to find the length of the diagonal. K-means algorithms are generaly easy to implement and have fast computational times, but have the disadvantage of pre-defining a number of stations.

COMPARISON OF MODELS

This section aims to enable fair comparison between the three models, while also keeping account for the practical applicability of the models. The first will be done via criteria for modeling ability discussed with Shell, the latter will be done by expectations of computational time and practical issues. A final verdict based on both factors follows in the final paragraph of this section: modeling decision.

Modeling criteria
Appendix E discusses ranked modeling criteria that were output from a discussion with Shell. These criteria are different than the performance measures as they are targeted at modeling abilities. For example, the ability of a model to locate the stations at any point or on a given network is included. In this line, the notion is made that the model should enable locating a station at any point. While it may seem counterintuitive to steer away from highways, coverage may be more optimal in some other way: proper modeling should serve the answer. This criteria of positioning the stations is ranked second. The ability to minimize the driving distance of trucks is considered the most important. This is to create initial customer traction. Being able to model the capacity of a station and to minimize the number of sites are rated lower. Table 6.1 displays the results of this ranking in the first column.

Table 6.1: Criteria for model comparison for Station Location Modeling - objective function is highlighted

Models	Flow interception	Clustering	FLP	?
Criteria / ranking	**NC-FRLM**	**K-means**	**ConFL**	**Aim of research**
1. driving dist. trucks	constrained	min.	min.	min.
2. position of stations	on road network	any point	any point	any point
3. capacity per station	capacitated	variable	variable	semi-capacitated
4. # sites	min.	fixed	min.	-

The columns next to the rankings show how each model handles the criteria. The objective function of each model is highlighted in yellow. The aim for modeling abilities for this research is displayed in the last column. Further explanation of the aim of this research is present in Appendix E. The table shows that the modeling abilities of k-means and ConFL are most consistent with what would be ideal, as determined by Shell. This coincides with the scores of the performance measures in figure 6.3a. The driving distance for trucks is ranked highest and drives the demand performance measures. While the criteria ranked lowest have their origin in supply-side performance measures. Nevertheless, all three models are evaluated based on their practical application to determine what is workable.

Practical application
In this paragraph, the practical application of the three models and the criteria is discussed. It is important to find a model that not only fits well theoretically, but is also practically aligned with the objectives of the thesis.

NC-FRLM
To draw up predictions on the practicability of the NC-FRLM, Daniel Speth, creator of the NC-FRLM is contacted [10]. Speth indicates that this type of optimization is too complex on this level of aggregation and

might not even be necessary. He flags that optimization often has little in common with reality, and that adding constraints will get closer, but also rapidly increase computational time. An estimate of computing power would require a computer with at least 30 GB of RAM (D. Speth, personal communication, March 23, 2022).

ConFLP
There are very few works on ConFLP, making it hard to draw conclusions on its practicability. However, it is known that the set of efficient solutions is seldom directly useful in practice. This is mainly because it usually is too large and contains many unsatisfactory solutions [100]. In practice, a highly precise description of the actual goal of the problem is needed. In fact, these should often be described by a globalizing or utility function that combines all the individual distances into a single global value.

K-means algorithms
As already noted, k-means algorithms are easy to implement and have fast computation times when dealing with a large dataset. However, it is often difficult to predict the correct number of clusters (K-value) for different types of datasets. In the case of this thesis, k-means algorithms would be used to distribute demand across gas stations: the clustering would not aim to identify x number of clusters. This problem is thus less relevant, especially taking into account that, according to the ranking of criteria, the number of locations is the least important variable.

MODELING DECISION
Based on the assessment of the practical application, NC-FRLM does not seem to be the right model for this study. Models for the location of gas stations process data based on truck flows. For this study, at a European aggregation level, this would mean that the model would have to process an extremely large number of flows. Doing so would require significant computational power, which would be out of scope of this research. On top of that, based on the criteria, the NC-FRLM scores lowest on allignment of goals of the model.

This leaves k-means algorithms and ConFL models, which score practically the same based on the criteria. Both have the objective of minimizing the driving distances of the trucks, assign the stations to any point, and have a variable capacity per station. The main difference lies in the fact that k-means works with a fixed number of locations. However, a subtle but important difference lies within that characteristic which has influence on capacity per station and the functioning of the model for this research in specific. The ConFLP determines the number of stations based on an optimization between costs of the trucks to drive towards the HRS and the costs to build an extra HRS: hence the dual objective function. If one of these two costs is not properly defined, the results can be highly skewed, leading to a highly unrealistic model. Far too many stations could be built because the cost for trucks is estimated to be relatively too high, or far too few stations could be built if the cost to build one station is estimated to be too high. This reflects in the previously touched upon general practicability of the ConFLP where was signalled that "in practice, a highly precise description of the actual goal of the problem is needed." As for hydrogen stations, there are still many uncertainties in terms of costs and they are also sensitive to changes over time; relying on ConFL models is therefore not recommended.

In contrast to the relatively unknown inputs to ConFL models, the biggest question with k-means lies in the initial number of stations. However, there are several predictions, even by country, about the number of stations that will open. Moreover, the number could also be aligned to fit the overall demand for hydrogen. **K-means algorithms** are therefore left for application. Use of such a model would ensure demand optimal station locations as is the goal of this step.

6.3. NETWORK SYSTEM DESIGN
After modeling the station locations, the next step of the optimization is the network system design. The demand-optimal locations, as determined by the k-means model, will be connected to the backbone (source of supply). This connection should be done in optimal fashion to create an optimal network system design. The connections to the backbone represent pipes which will have dimensions in length and capacity. Based on these dimensions, calculations of CAPEX of supply would be able. Figure 6.4a shows how the network system design should influence the performance measures. Figure 6.4b displays how this influences the balance in the two-step optimization.

Performance measure	Focus NSD
1. CAPEX	+++ (supply)
2. Utilization	+/-
3. Coverage	+/-
4. Distance	+/-

(a) Performance measures Network System Design

(b) Performance focus Network System Design

Figure 6.4: Performance focus Station Location Modeling

With respect to the design of new energy networks, Heijnen *et al.* conducted a thorough literature review in the areas of operations research and systems engineering. In general, they write on three approaches to solve network system design problems, which are discussed below, along with their (dis-)advantages.

Mixed Integer (Non-)Linear Program (MI(N)LP)

Mixed Integener (Non-)Linear Program studies are often used to analyse network systems [103–106]. The MILP requires an objective function and a set of constraints to arrive at an optimal solution. This has the advantage that the optimization problems can be formulated in a clear manner, but the disadvantage that the adaptability is rigid. Several spatially-explicit infrastructure models have been developed. Most of them are not capable of constructing a regional pipeline network to connect multiple production facilities and demand locations, as would be required. This is however, incorporated for the modeling of a hydrogen network by Johnson and Ogden (2012).

Agent-based models

Using ABM to find (sub)optimal network designs is a bottom up approach. These models rely on ant colony optimization (ACO) [108, 109] or particle swarm optimization (PSO) [110, 111]. The agent-based method is easy to apply, as the NetLogo implementation is quite intuitive [112]. The decision rules in the model can be changed at will to address other problem characteristics, making the model flexible.

Heuristics and algorithms from graph theory and geometry

Heuristics and algorithms from graph theory and geometry have long been added to the literature. The relatively simplistic structure of this theory has allowed a large toolbox to be built up within the field. In comparison with ABM, they hardly differ in the quality of solution and computational performance [112]. However, since these algorithms are specifically designed to solve network design problems, Graph theory is able to provide a balance between the flexibility of agent-based models and the rigidity of MILP [8].

6.3.1. FIT FOR THIS THESIS

As indicated, graph theory is able to provide a balance between the flexibility of agent-based models and the rigidity of MILP. For this thesis, this balance is welcome since the network design must be built on previous modeling steps. Processing the results of the freight demand model and the subsequent station location model requires flexibility, as a large data set is being analyzed. Moreover, graph theory is specifically designed to analyze networks, while MILP- and agent-based models have other primary applications [112]. Hence, use of such models would require more effort to be applied for network system design. Therefore, the goal is to implement a scientifically correct hydrogen backbone, based on graph theory, and link the HRS sites to it. However, as with the previous chapters, the choice of a theory may also depend heavily on available data.

DATA SEARCH

Since the hydrogen economy is still in its infancy, no (model of an) actual hydrogen pipeline network is present. Therefore, a decision on station locations, taking into account the distance from such a network, should be based on a predictive study. To this end, a novel methodology combining graph theory and exploratory modelling and analysis (EMA) is used by Huisman (2021). In his work, he mapped out the growth of a hydrogen network in three stages. In line with by H2Accelerations expectations, the network will grow by various small connected components that span initial adoption centers. These centers are situated in Germany, Italy, Belgium and the Netherlands. The timeline is comparable to the plans of H2Acceleration, as Huisman considers first efforts to be present in 2030. In steps of a decade, these components are modeled to grow in size, eventually connecting a large connected network in 2050 with Germany and Italy at its core. The network is displayed in figure 6.5, indicating the date of development and the capacity by the color and the width of the pipeline respectively. One can see that it extends from France in the west to parts of the Balkans in the east. It should be noted that a large number of uncertainties in terms of demand, supply and geopolitics are taken into account.

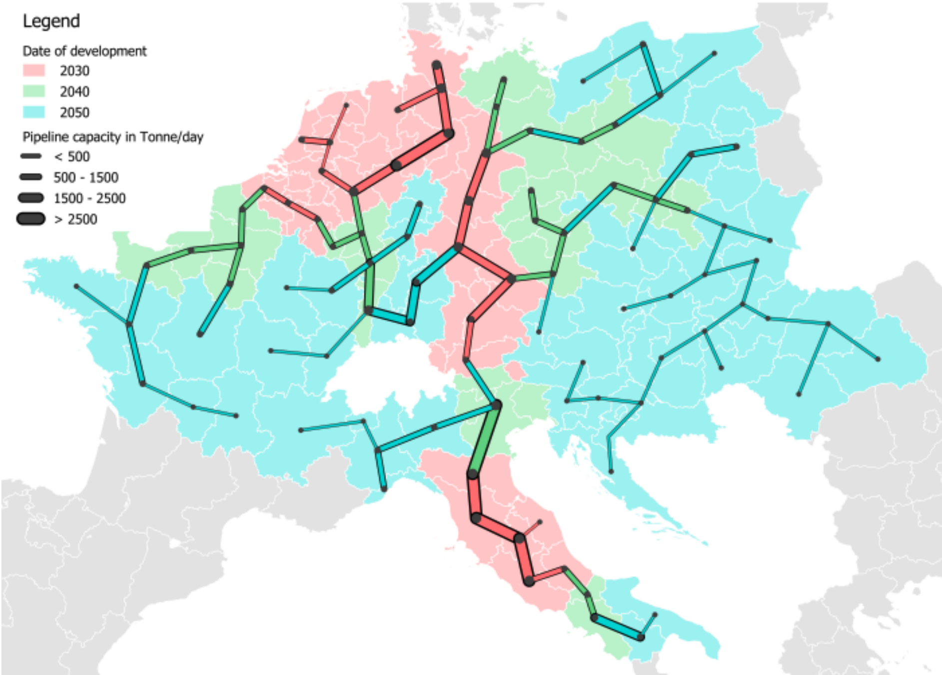


Figure 6.5: Growth of robust hydrogen backbone between 2030 and 2050 as determined by Huisman (2021)

To fit the research objective, the network design method must be able to process the costs of supplying a refuelling station via hydrogen backbone, over time. The design of the European hydrogen network by Huisman (2021) is built on graph theory, which ensures being able to process such costs. The availability of this backbone, with data available of growing clusters over time, reconfirms the choice for **heuristics and algorithms within graph theory** as the desired method.

6.4. TWO-STEP OPTIMIZATION

In this final section of the theoretical framework, the two-step optimization is highlighted. As stated in the introduction this final step is called "two-step optimization" for the sake of reference. In reality the entire process of the previous two sections combined with this section is the two-step optimization. Nevertheless, this step brings them both together, making the name appropriate.

The first two steps were the station location modeling and the network system design. This results in a demand-based optimal topology of stations, supply-optimally connected to a hydrogen backbone. This network is already optimal in its own way, however costs of supply are not yet taken into account in the locations of the stations. This is because the network system design optimizes the costs of supply without moving the stations itself. In this step, the costs of piping that are a result of the network system design are used to create a new topology. This should have a balancing function on the performance measures as is sketched in figure 6.6.

Figure 6.6: Search for balance in theoretical framework with performance measures

Relocating the stations should be done by finding station locations that balance out the costs of supply with costs of demand. This means that the cost of hydrogen to get there by pipeline is as high as the cost to get to the HRS for all trucks. This balance can be sought by drawing up a **balancing cost function** with the aim to minimize total costs that are the sum of costs of supply and demand. Chapter 11 will search for such a function and apply it on the networks.

6.5. CONCLUDING REMARKS

The theoretical framework was intended to provide a skeletal structure within which appropriate methods would surface. A two-step optimization proved to function well as a means to balance the costs of supply and demand. By first getting demand-optimal locations with **station location modeling** and optimizing their supply with a **network system design**, an initial network is created. This can be further optimized in the **two-step optimization** by finding station locations that balance out costs of supply and demand. The visualisation of the theories along with their methods is displayed in figure 6.7

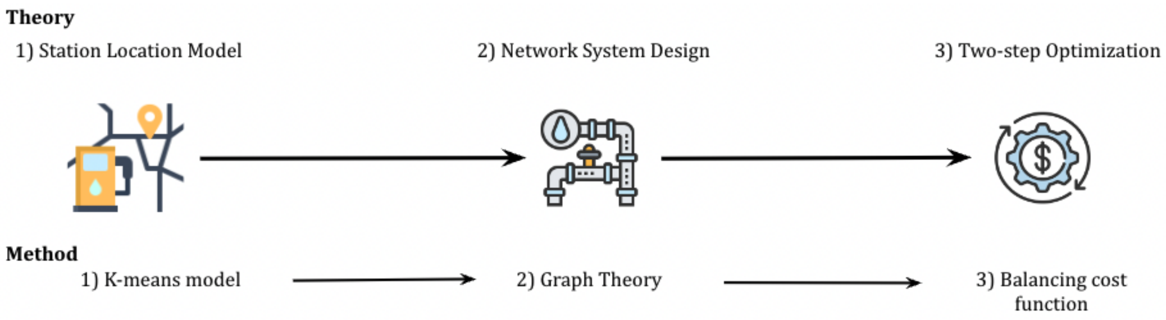


Figure 6.7: Theoretical framework with methodologies

In terms of methodologies, choices were made that fit in with the two-step optimization goal of balance. A **k-means algorithm** will bring the refueling stations into optimal place in terms of demand. **Heuristics and algorithms in graph theory** will ensure optimal costs of supply to those station locations. Finally, use of a **balancing cost function** will ensure inclusion of the costs of supply in the station locations. Note that each step requires output of the previous model to function, where the station location model requires the output of the freight demand model. This is summarized with the rest of the inputs and outputs in figure 6.8

1a. Input:
Freight demand model
- dispersed hydrogen demand

Number of HRS
- average station size

1b. Output:
Optimal demand-based topology
- HRS locations and demand

2a. Input:
Demand-based topology
Hydrogen backbone model
- Costs of pipelines

2b. Output:
Optimal network system
- HRS + piping system
- Costs of supply

3a. Input:
Optimal network system
- Costs of supply
- Costs of demand

3b. Output:
Optimized station topology

Figure 6.8: Input and output of methodologies within Theoretical Framework

Chapter 7 and 8 elaborate on the basis of the model, as the freight demand modeling will be performed and the analysis of the regulatory basis will be discussed. The chapters thereafter (9, 10, and 11) will function as the methodology and initial application of the station location model, the network system design and the two-step optimization just highlighted. These three chapters present and construct the core model of this thesis.

7

Freight Demand Modeling

This section presents the methodology and initial application for freight demand modeling. Section 7.1 lays out the methodology proposed by this research, building on the foundation laid in the basis for modeling (chapter 5). The methodology is built from that of an external source and it is explained how it was adapted. The sub-question that this section is looking to answer is identified in the table below. By calculating European hydrogen demand for HDTs and distributing them over a highway network, initial results will come to light. However, more importantly a basis for the rest of the research is established. This will be presented in section 7.2.

Table 7.1: Research Goals of Freight Demand Modeling

Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion addressed in this section	**Research Method**
What is the expected European hydrogen demand from heavy-duty trucks over time, distributed over a highway network?	Economic input-output models

7.1. Methodology

From the theoretical framework, the economic input-output method remained as the appropriate method for this thesis. This was mainly due to the availability of economic data on NUTS3 level. It was found that data for NUTS3 regions is suitable for distinguishing sub-national regions at a scale small enough to generate interregional (and international) truck flows. Eurostat has published a database of tonne-kilometres (TKM) for freight transport at national and EU level. They include breakdowns into domestic transport, international transport and cabotage [113]. Based on these data, economic input-output models can synthetically simulate flows of trucks. Figure 12.18 shows Belgium divided into its NUTS3 regions for reference.

7.1.1. Data Processing

In order to create truck flows, economic activity between the NUTS 3 regions must be processed so that estimates of freight transport from origin to destination can be derived. The demand, from each origin to destination, can then be plotted over a network of highways using a shortest path algorithm. This way, a synthetic flow of European trucks will cover the European highway network as is asked for by the research objective. One source is found that has processed the NUTS3 data in such manner [10]. Their resulting dataset is publicly available and used for this thesis. An explanation of the creation of their dataset is discussed in the text below, broken down into the methodological steps that were taken.

Figure 7.1: Example of NUTS 3 level: Belgium [9]

Model development

For data on origins and destinations, the results of the European Transport Policy Information System (ETIS) are used. The ETISplus 2010 project provides a very comprehensive survey of European transport [114]. This ETISplus dataset contains an OD matrix for road freight transport, which is used as the basis for the model. The generation of this OD matrix used Eurostat's transport data tables at the NUTS-3 level (discussed earlier), as well as national databases. But, as the name implies, the OD matrix stops at 2010. The figures up to 2010 have to be scaled up to 2019.

Growth rate up to 2019

To calculate the flows for 2019, the same Eurostat tables (but for 2019)[113] are used to calculate national growth rates. Annual (inter)national transports are broken down by the respective NUTS-3 regions in which the goods were loaded and unloaded. It must be noted that data availability is inconsistent and incomplete for some countries and years. To cope with this lack of data, a country-specific growth factor is calculated based on aggregated national and international transport flows [10]. It corresponds to the relative change in transport volume of the individual countries compared to 2010. As the average growth rates of exports (3.7%) and imports (3.64%) are almost equal, the country-specific export growth factor is used to scale all international transport flows.

Growth rate up to 2030

After a growth factor is used to get numbers for 2019, traffic flows are calculated for the year 2030. As there is no single growth value that is agreed upon in literature, this model assumes the same growth rate between 2019 and 2030 as there was beteen 2010 and 2019.

Conversion to trucks

To convert the freight volume into trucks on the road, loading factors are required. The European Commission reported an average loading factor of 13.6 tonnes in 2010 and this remained practically constant in the years after. Therefore, the model assumes a loading factor of 13.6 tonnes for 2010, 2019 and 2030 [10]. However, this average loading factor refers to transport of loaded trucks where the route calculation refers to transported freight volumes. This means that empty runs are not considered. In 2018, the average percentage of empty runs was 20% for HDTs. To compensate, the number of trucks are scaled upwards by a factor of 1.25. This number stems from the difference between the synthetic data and real life truck traffic counts from Germany, as will be discussed in the uncertainties.

Road network development
To run the flow between origins and destinations, a network must be present. In the model, a highway network was extracted from the ETISplus road network and manually updated with the current E-road network [115]. To reduce complexity of using all network elements in the dataset, the network is reduced to road sections that are part of a highway or the international E-road network. On top of that, important ferry connections are also added manually. The speed of a truck is assumed to be 80 km/h.

Routing
To assign the calculated transport flows over the road network, transport routes need to be determined. The first step is to assign every NUTS3 region to a network node that is closest to the middle point of the region. This way, start and ending points of the transport routes are determined. To incorporate routes from O to D, Dijkstra's algorithm is used. This shortest path algorithm computes the route with the minimum distance for each OD pair.

The results of the actions listed in the previous sections are shown in Figure 7.2

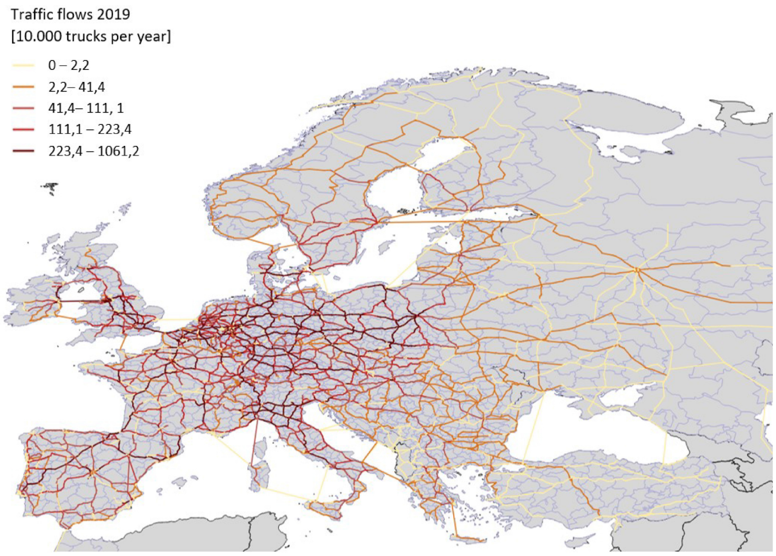


Figure 7.2: Traffic flows 2019 as modeled by Speth *et al.* (2022) [9]

RESULTING DATASETS

The synthetically generated road freight traffic volume for each road section provide an extremely relevant basis for the design of the hydrogen refuelling station network. Speth *et al.* (2022) created four individual data sets that are publicly accessible:

1. the truck traffic flow data
2. an overview of the included NUTS-3 regions
3. a list of the network nodes in the underlying road network
4. a list of the network edges in the underlying road network.

The first dataset provides truck traffic flows between each NUTS-3 region. In total, it considers 1,514,573 directed transport flows between 1630 different origins and 1667 destinations.

The overview of NUTS-3 regions contains data on the regions themselves (ID, name, country) and the location of their centers. A nearest point in the E-road network is also assigned.

The third and fourth datasets describe the unverlying E-road network, which consists of 17,435 nodes and 18,447 edges. Per edge, truck traffic volumes are provided for 2019 and 2030.

7.1.2. TAILORING DATASET FOR THESIS

The datasets contain truck flows, which can be used to determine station locations. However, to enable HRS location modeling by k-means (section 9) the truck flows need to be represented as data points. Moreover, the data currently considers "regular" diesel trucks whereas a demand for hydrogen is to be modeled. Therefore, this section is dedicated to tailoring the data to fit the needs of this thesis.

Truck flow to datapoints

To transform the given flows into datapoints, a python script is written. In this script, the geographical middle of the edges is computed. This is possible because each edge is a connection between two nodes, whose latitude and longitude are known. The geographical mean of these two nodes represents the middle of the edge. Figure 7.3 illustrates this process. As there were 18,477 edges, the conversion results in an equal amount of data points with a latitude and a longitude. This leads to major computational gains, as 1,514,573 flows are translated into 18,477 points.

When converting truck flows into individual data points, it is important that the information represented by the flow not be diluted or double counted. The flow of trucks from origin to destination is routed over the underlying highway network of nodes and edges. Thus, the 18,447 edges of the network are each assigned a different flow of trucks. These numbers, trucks per year, are known per edge for 2019 and 2030. This data is assigned to each data point, causing the total flow of trucks to be represented by the datapoints. However, simply appointing the respective flow on an edge to the corresponding datapoint would imply doublecounting. Therefore, the number of trucks are recalculated to vehicle kilometers.

Trucks to vehicle kilometer

As briefly introduced, each edge had a corresponding traffic flow for 2019 and 2030. The inherent problem is that when transferring this data to points, the core difference between a flow and a point gets in the way: a flow is a connected network, where datapoints stand alone. When looking at trucks as a flow, a truck traveling from O to D passes many edges but it does not get double counted. When individual data points are considered as a network, the total truck traffic flow would be too high due to double counting. Each truck would be counted at each data point it passes through, skewing the data. To account for this, the data is modified. Instead of assigning a number of trucks to each data point, vehicle kilometers are calculated for each data point as is common in transportation studies. This is done by multiplying the distance (length) of each edge by the number of trucks that pass over it per year. This way, double counting is prevented. All the above steps are shown in the figure below.

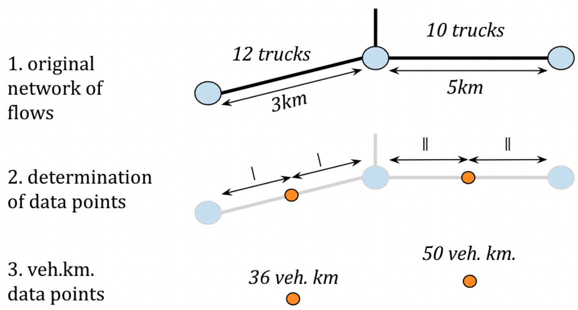


Figure 7.3: Explanation of creation of datapoints

HDT traffic to FC-HDT traffic

As the dataset of Speth *et al.* (2022) considers any type of heavy-duty vehicle, the numbers must be adjusted to fit the objective of modeling fuel-cell heavy-duty trucks. To incorporate this, the market penetration rate of FC-HDT (as identified in section 4.5) is used. The total number of vehicle kilometers on each datapoint are thus multiplied for each decade by the respective market penetration: 5% in 2030, 15% in 2040 and 30% in 2050.

7.2. INITIAL APPLICATION

The 18,447 datapoints can be plotted in Python, based on their lattitudes and longtitudes. To visualize the "flow", each datapoint can be colored accordingly to the weight of vehicle kilometers per datapoint. As these vehicle kilometers wildly differ, a logarithmic scale is used. Edges that have a flow 0 are removed:

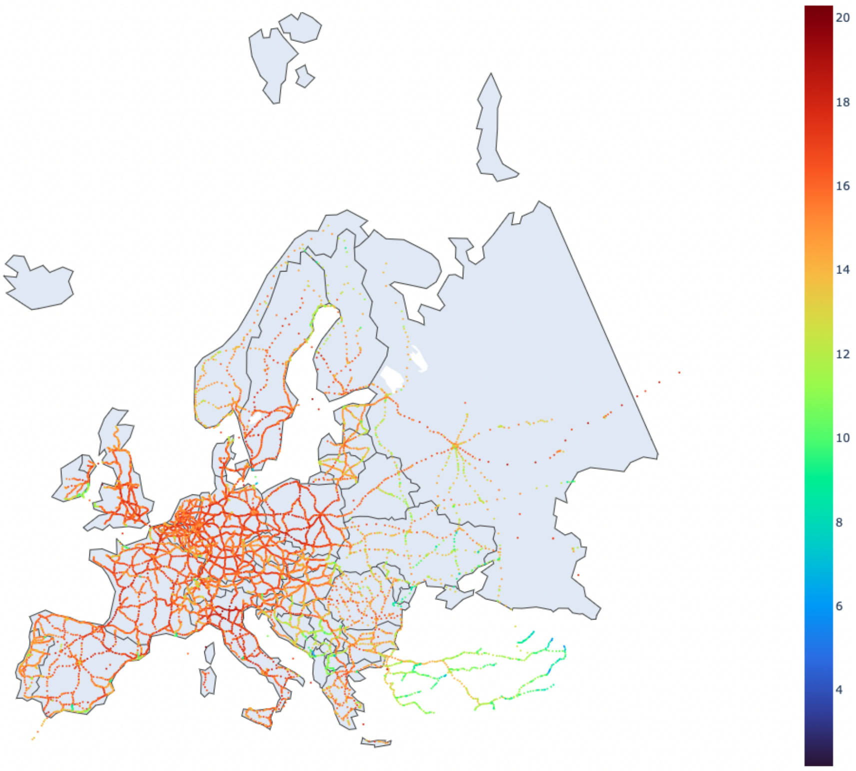


Figure 7.4: Plotted traffic flow of 2030 displayed as datapoints

Note that the map displayed in figure 7.4 is the traffic flow plotted for 2030. The same figures and calculations are performed for 2040 and 2050, as the intention is to model demand over time. The growth rate that was applied on 2019 to get 2030 numbers will again be used to get numbers for 2040 and 2050.

RECALCULATION TO DEMAND FOR HYDROGEN

To answer the third research question and get to a distributed demand for hydrogen, the number of vehicle kilometers can, combined with the market penetration rate (A.2) be recalculated to local demand for kilo's of hydrogen. Based on section A.2, it assumed a fuel-cell heavy-duty truck can drive 1440 km on 85 kg of hydrogen (at 700 bar). Therefore a FC-HDT can drive 17 km (1440/85) on 1 kg of hydrogen. Hence, 17 vehicle kilometer on a datapoints represent demand for 1 kg of hydrogen.

7.3. CONCLUDING REMARKS

The intention of the freight demand model was to create a basis for further modeling. To account for the research objective, and avoid diesel data biases, results equal to truck traffic counts were aimed for. A fitting dataset was found which represents synthetic truck flows throughout Europe. By adapting this data set from flows to datapoints, vehicle kilometers are represented as points which enables modeling in light of the performance measures. The dispersed demand for hydrogen can be further calculated by the efficiency of the fuel cell heavy-duty trucks. The station location model will make use of this data in chapter 9. Next to that, it will be used to analyze the regulations of AFIR in chapter 8.

8

Regulatory Basis

This section presents the methodology for analyzing the regulatory basis. As discussed in section 5, the regulatory constraints for a network of refueling stations will not function as a basis for modeling in this study. However, they will be analyzed to provide insight into the ongoing regulatory negotiations. Section 8.1 describes the methodology and pre-processing steps to get there. In chapter 12, the regulations of AFIR will be analyzed accordingly. The sub-question that this section supports to answer is identified in the table below.

Table 8.1: Research Goals of Freight Demand Modeling

Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion addressed in this section	**Research Method**
What are the current plans and institutions for hydrogen refueling stations and how do they influence the network design?	p-median modeling

As a reminder, the constraints that stem from regulation are repeated:

- a hydrogen refueling station should be present (700 bar, 2t/day) every 150 km along TEN-T corridors
- a hydrogen refueling station should be present in every urban node

Both these constraints stem from the proposal of Alternative Fuel Infrastructure Regulation (AFIR). The focus of analysis of these constraints should be in light of the previously mentioned performance measures of a network of refueling stations. Of the four measures, only two are relevant for analysis: over- or underutilization and network coverage. The other two (CAPEX and average distance to trucks) are not relevant for analysis. This is because both are already present in the regulation: distance in the 150 km remark and CAPEX because a fixed number of stations and a fixed size are proposed.

8.1. Methodology

This section discusses the methodologies to test the AFIR network on two performance measures. First, a potential analysis for under- or overutilization is discussed followed by a methodology for the coverage of demand. For both analyses, the distribution of hypothetical demand of chapter 7 is used. But first, the AFIR station locations must be usable for analysis.

Pre-process

In order to analyze the regulation, hydrogen refueling stations are placed at the coordinates of the list of urban nodes from Annex II part 2 of the TEN-T Guidelines. The latitude and longitude are obtained from a latitude and longitude finder in python.

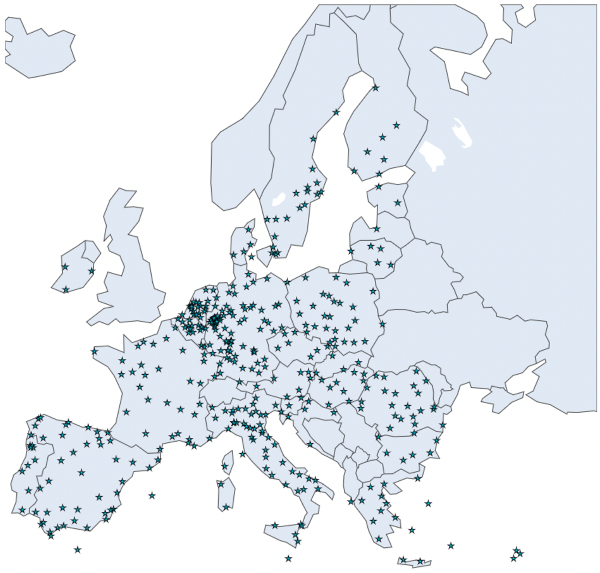


Figure 8.1: Initial AFIR locations - urban nodes

Secondly, it is examined for which part of the TEN-T corridors the 150 km maximum distance limitation is then not yet sufficient. By considering only data points that fall within a 150 km range of each "urban node station", almost the entire network (and no more) as shown in the figure 8.2 is considered. A modeling step in Python is written to find out where an HRS would be lacking. Consequently, some HRS are manually added to attain to the regulation of having an HRS every 150 km, these are highlighted green in figure 8.3. This set of locations, in combination with the freight demand model, enables further analysis.

Figure 8.2: AFIR network as given by TENtec

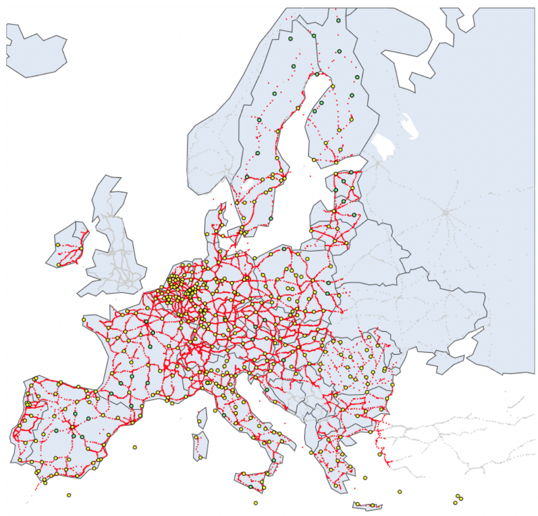


Figure 8.3: AFIR network as modeled

AFIR analysis: utilization

To analyze under- and overutilization of the suggested station locations the AFIR stations are placed on the same plain as the results of the freight demand model. Thereafter, the demand from each node can be assigned to the station closest to it. By doing so, the potential uptake of each AFIR location will become clear. The stations are prescribed to have a capacity of two tonnes per day. Over- or under utilization can be calculated accordingly.

AFIR analysis: coverage of demand
The other aspect to be analyzed is the coverage of the network. By capacitating the total demand that can be appointed to an HRS to 2 tonnes per day, the coverage of each station can be analyzed. Since the location of demand points, as well as the locations of the HRS are known, a p-median model can be used for this purpose. More specifically, a capacitated p-median model would suffice the goals of thesis.

In the theoretical framework, the p-median model was already shortly touched upon. Being a discrete location model, the fit was not there. For the station location model, initial locations were supposed to be determined by demand and not predetermined. The input for a p-median model is a given finite set. This is exactly what the AFIR proposal provides. The model searches for the closest demand points in its proximity and adds them to a list of "served" demand points until the stations capacity is reached [116]. The capacity added to the HRS is equal to the vehicle kilometers on that datapoint. Note that, if demand of a datapoint in proximity is too large, it cannot be appended.

8.2. CONCLUDING REMARKS

The intent of this section was to provide a methodological basis for analyzing the regulations affecting hydrogen refueling station locations. As such, the station locations determined by AFIR will be analyzed for two facets: utilization and coverage. This will be done by placing the locations in the same model plane as the freight demand model and then assigning demand to the stations. For coverage, a fixed capacity of two tonnes per day will be assumed to check how much of the demand will be covered. For utilization, no capacity is assumed and the demand closest to a station is allocated to respective HRS. In this way, potential utilization and a minimum utilization rate can be determined. This analysis will be performed later in Chapter 12. The coming three chapters will present the methodology and initial output of the main focus of this research: the two-step optimization.

9

STATION LOCATION MODELING

This chapter presents the methodology and initial application of the station location model. Section 9.1 lays out the methodology for the optimization proposed by this research, which builds on the theoretical foundation of chapter 6. The sub-question that this section is looking to answer is identified in the table below. Distributing hydrogen refueling stations in a way that optimally meets demand reveals results and provides a basis for further optimization. Initial results will be presented in section 9.2. The chapter concludes with a summary of what was presented, combined with an overview of where we are in two-step optimization.

Table 9.1: Research Goals of Station Location Modeling

Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion addressed in this section	**Research Method**
How are hydrogen refueling stations distributed over Europe in order to optimize coverage of European FC-HDT demand?	K-means algorithm

9.1. METHODOLOGY

As established in the theoretical framework, station locations will be modeled via k-means clustering. Because there is no regulatory basis, the locations and sizes of stations will be modeled purely on the basis of demand. Note that also no current refueling sites are considered. This leaves out specific forms of bias, as laid out in the research objective. The results will be station locations that are optimal in terms of demand.

9.1.1. IMPLEMENTATION OF K-MEANS ALGORITHMS

K-Means algorithms are a class of unsupervised learning algorithms used to distinguish clusters of data in a given dataset. The advantage of an unsupervised learning algorithm is that it does not require labeled data in order to find patterns in the dataset [101]. The strict rule, and goal, of clustering is that datapoints are divided in such manner that they are more similar to points within their cluster than those outside of it. The most common way to compare similarity of datapoints is by measuring distance. This method of clustering also fits the purpose of this thesis, as datapoints are to be clustered by their distance to an HRS. Such similarity can be found by using Pythagorean theorem to find the length of the diagonal: the Euclidean distance. Other distances can also used (e.g. Manhattan distance) but for this thesis Euclidean distance is the right fit. This process of clustering leads to an optimal distribution of centroids (HRS) to cover datapoints (demand nodes).

K-Means is a form of partitional clustering, implying that it takes into account a predefined number of clusters and will divide the data into exact that many partitions. This entails the characteristics that no cluster can be empty and that each datapoint can belong to only one cluster. Moreover, it should be noted that K-Means algorithms are non-deterministic and thus may end up with different clusters each time they are executed.

K-Means process

To clarify how a k-means algorithm works, pseudocode and an example are useful. The pseudocode describes the k-means process step by step, followed by a description that refers to the illustrative example:

1. Specify number of clusters (k)
2. Initialize k centroids (random or ++)
3. DO
   a. assign each datapoint to closest centroid
   b. update centroid location by computing mean of each cluster
4. END IF
   centroid location is not changing

For the example given below, a Guassian mixture is used as data points as can be seen in 9.1a. The first step of the algorithm is specifying the predefined number of centroids. The centroids are the centers by which the data will be clustered and are representative for the HRS in this thesis. For this example, 3 centroids are chosen (k = 3). The second step is determining the initial location of the centroids. Two types of initialization are most commonly used for k-means: random and plusplus initialization. Random initialization places the initial centroids randomly across the plain while plusplus places each centroids sequentially, taking into account distance to the previously placed centroids. Plusplus initialization often gets the better results. However, randomly chosen centroids better illustrate the operation of the algorithm and are therefore used for the example. Figure 9.1b shows pseudocode steps 2 and 3a. The (colored) centroids are placed randomly and then the data points are clustered (colored) based on which of the 3 centroids is closest.

(a) 0. Gaussian mixture of datapoints
1. k = 3 (centroids)

(b) 2. Random initialization of centroids
3a. assign datapoints to cluster

Figure 9.1: K-Means clustering (random initialization) step 1, 2 and 3a

Next, the locations of the centroids are updated by taking the mean of the locations of the datapoints in their respective cluster (step 3b). This is displayed in figure 9.2a, and shows how the centroids naturally become a cluster center. This process repeats itself until the centroid locations do not change, signalling that no shift of datapoints to another cluster has taken place (step 4). As figure 9.2b shows, the centroids now represent a center of datapoints that minimalises distance from the assigned points.

(a) 3b. update centroid location mean of cluster

(b) 4. Centroid location is not changing

Figure 9.2: K-Means clustering (random initialization) step 3a and 4

9.1.2. INITIAL NUMBER OF STATIONS (K)

The performance of K-means clustering is usually dependent on the initial number of centroids [117]. There are several methods to determine the number of centroids that will make a valuable set of clusters. For each dataset and clustering purpose, it might differ what fits the needs best. The goal of this thesis is to design a realistic optimized network, able to cover the demand of FC-HDTs. The clustering will serve to optimally locate an x number of stations that can serve that demand. In this section, the most popular method for picking k will be discussed followed by a tailored method.

ELBOW METHOD

The most popular method for determining the initial number of centroids is the "elbow method" [118]. The elbow method is based on running the K-means algorithm multiple times, increasing the number of clusters by one iteration each time. The measurement of performance of the number of clusters is the level of inertia. The target is to minimize this inertia (or "loss"), hence seeking to diversify the clusters as much as possible. If each datapoint is given its own cluster center, this is maximized. However, this would not add value and so a balance is sought. For this case, it would be measured in distance to a cluster center. This inertia will be plotted for all runs against the number of clusters, on a line graph. The curve of inertia starts to bend at a point deemed a reasonable trade-off between error and the number of clusters. This point is called the elbow point and is visually picked as a value for k.

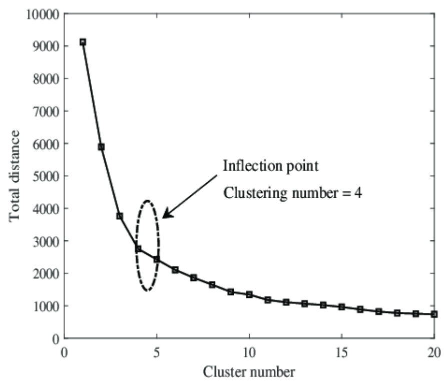


Figure 9.3: Example of the elbow method [11]

Application

To see if the elbow method is useful for this thesis, it is applied to the dataset. As shown below, the elbow point of the dataset forms at 10 clusters. This suggests that adding more stations to the network will yield relatively little improvement in coverage. However, this needs to be put in perspective.

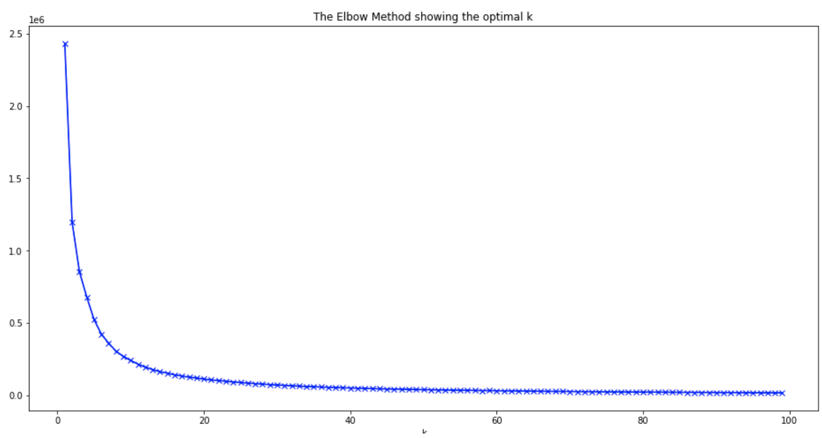


Figure 9.4: Elbow method for clustering 2030 dataset

As shown in the graph, distance loss is measured in very large numbers due to the size of the network and the long distances of Europe. This entails that, after the creation of 10 stations the distance gets reduced by a relatively small amount compared to the addition of the first 10 stations. However, the absolute gain is still very large. The goal of this thesis is to create a network of refueling stations that covers Europe. In order to create a coverage network, the initial distance from any truck to an HRS should not be too great. The marginal gain that is to be made by adding an HRS seems negligible when using the elbow method however the absolute value needs to be further reduced to create exceptable coverage. If demand is to be served in a propper manner, a different method for k determination must be sought.

TOTAL DEMAND COVERAGE

To meet the objectives of the study, the number of stations must be large enough to cover the European demand for hydrogen. Therefore, a tailor-made method can be developed to arrive at a number capable of doing so. This number of stations can be derived by assuming an average-sized station and dividing its capacity by the total demand. The average allocated demand to each station will then inherently add up to the total demand for hydrogen. Note that this does not necessarily mean that all stations will be of the same size.

As Shell plans to standardize near-future stations at 2 tonnes of hydrogen per day this is taken as a fitting average to determine the number of stations for 2030. This number is also in line with the EU prescription for HRS in AFIR (section 4.4.2). However, for 2040 onwards it is expected larger stations will be the standard [12]. Demand will grow and creating more smaller stations is less economically viable. For this reason, 6 tonnes per day stations are taken as the average for 2040 and 2050. Using numbers of the system description (Appendix A), a two tonnes per day HRS can provide for 34.000 vehicle kilometers per day (2000 kg * 17 veh. km). As the data covers vehicle kilometers per year, this is recalculated to 12.418.500 veh. km. per year (365.25 * 34.000). The number of FC-HDT kilometers is already calculated in section 7.1.2. Dividing these totals by the capacity of an average station gives the following number of stations:

Table 9.2: Number of stations per decade

	2030	**2040**	**2050**
# stations	964	1285	3527
# size (t/d)	2	6	6

As mentioned, this number of hydrogen refueling stations will act as input to the k-means model. The number could function optimally to cover demand across Europe.

9.1.3. WEIGHTED K-MEANS

The k-means process, as illustrated in section 9.1.1, enables optimal coverage of a network of datapoints (demand nodes). However, to account for the different numbers of demand each node represents, the k-means clustering must be weighted. The difference between weighted and unweighted k-means modeling lies in the recalculation of the centroid location. Whereas in unweighted modeling the location of the centroid is determined by the average of the locations of data points within the cluster, in weighted modeling this location is determined by the weighted average of these locations. Each data point gets assigned a specific weight, which would in the case of this thesis be the demand for hydrogen (or vehicle kilometers). To illustrate how this method works, an example comparing both forms of k-means models is given in figures 9.5 and 9.6:

(a) Unweighted datapoints

(b) Unweighted k-means clusters

Figure 9.5: Unweighted K-Means clustering example

Figure 9.5a shows a set of unweighted datapoints, which are clustered by unweighted k-means in figure 9.5a. To show the works of a weighted model, weights are added to datapoints with y>8 and x<-1 or y>4 and x>1. These datapoints are appointed a weight of 20 (as opposed to 1) and are highlighted in red in figure 9.6a. Weighted k-means is applied, which yields the outcome as displayed in figure 9.6b.

(a) Weighted datapoints (red = weight 20)

(b) Weighted k-means clustering

Figure 9.6: Weighted K-Means clustering example

Figure 9.6b shows how the weights influence the location of the centroid. As can be seen, the optimization is mainly local, as the new location is determined based on the weights of the datapoints within each cluster. However, the centroids of the green cluster also moves partially due to the movement of the yellow centroid. For the network of datapoints that represent the European highways, it is expected mainly local optimization will take place as datapoints lie close together and no clear clusters will form.

9.2. INITIAL APPLICATION

To create a basis on which recommendations for Shell can be made, the initial networks that result from the weighted k-means model are presented in this section. The datapoints of the highway networks and their respective vehicle kilometers are used as input for the algorithm. Results, in the same format as the examples are displayed in figure 9.7 for 2030.

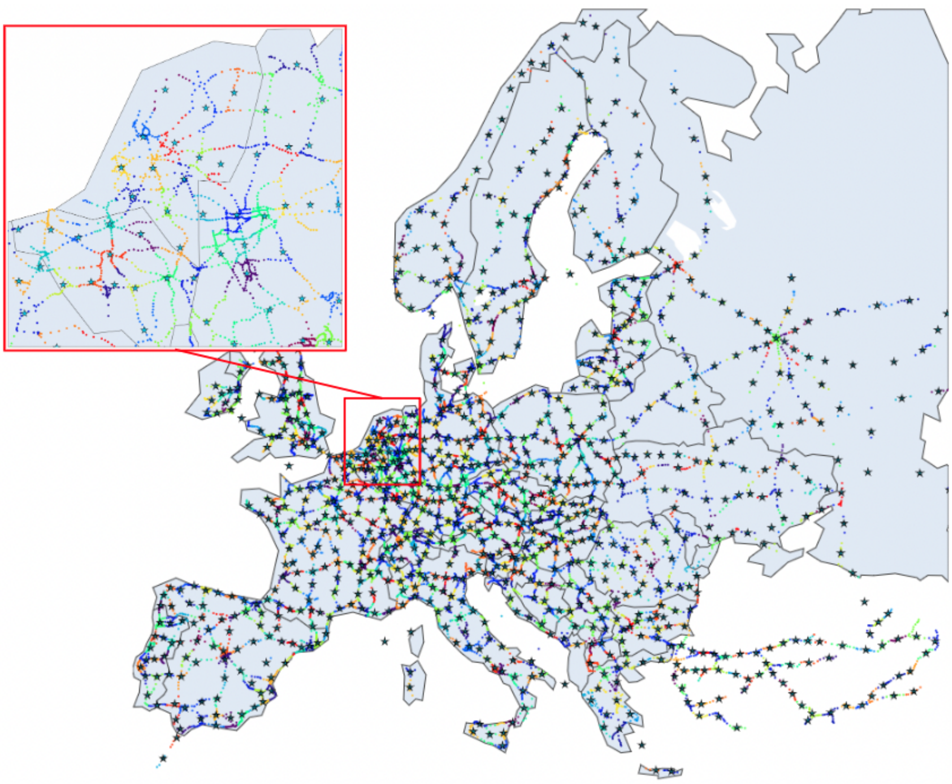


Figure 9.7: Weighted k-means application 2030 with zoomed-in section on The Netherlands and Belgium

The coloring shows the assignment of the data points to their respective HRS, represented by stars. As can be seen the network of datapoints is fully covered with small distances from each datapoint to an HRS. The influence of the weights seem to be mostly local, but have small continental effects as stations are more represented where demand is higher. Figures 9.8 and 9.9 show the same results but for 2040 and 2050. The figure for 2040 shows a very similar network as that of 2030, which makes sense as demand will grow by a factor of 3 as will the average station size. For 2050, the station size does not grow but demand increases with a factor of 2. Therefore, more stations are present creating an even better coverage of the network. These results are further analyzed in chapter 12.

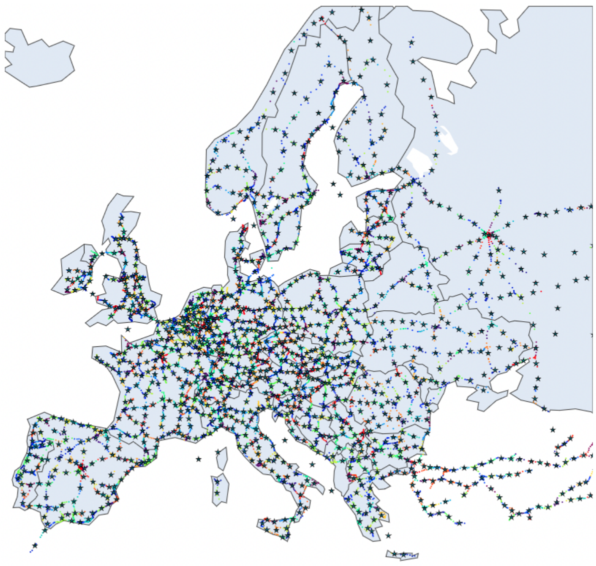


Figure 9.8: Weighted k-means application 2040

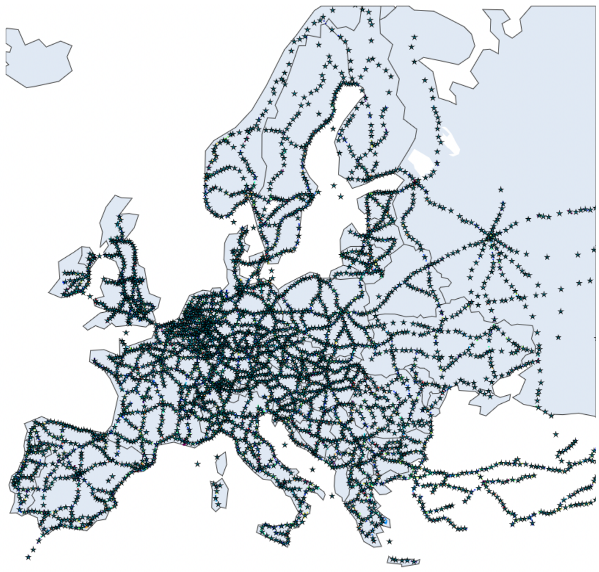


Figure 9.9: Weighted k-means application 2050

9.3. CONCLUDING REMARKS

This chapter attempts to present the methodology and initial application of the station location model of this study. The process of (weighted) k-means clustering is discussed, followed by the determination of the initial number of stations (k). In the second section, the weighted k-means model is applied to the freight demand model dataset, yielding initial results. These results are the demand-optimal station locations and represent the first step in the two-step optimization. As the figure below shows, this step will be input to the next step, the design of the network system.

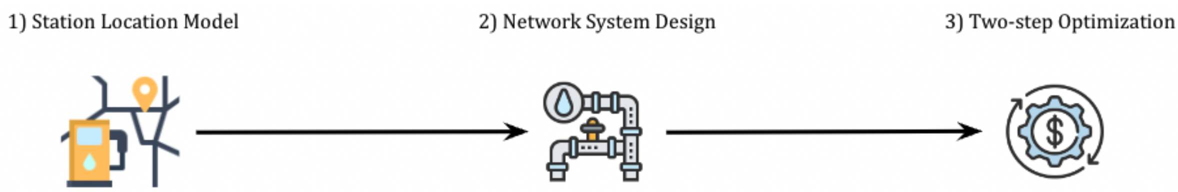


Figure 9.10: Skeletal structure theoretical framework

The next chapter (10) will use the output of this model as input, and build the network system design. The results of both chapters (9 and 10) will be used in the two-step optimization. All three of these will be further analyzed in section 12.

10

Network System Design

This section presents the methodology and initial application of the network system design. Section 10.1 lays out the methodology to connect the set of HRS to a hydrogen backbone to create a network system. The methodology builds on the theoretical foundation of chapter 6. The sub-question that this section supports to answer is identified in the table below. The creation of a network system will be a first step in analysis of the influence of a hydrogen backbone on optimal station locations. Initial results will be presented in section 10.3 and will be further analyzed in chapter 12.

Table 10.1: Research Goals of Network System Design

Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion addressed in this section	**Research Method**
What is the influence of hydrogen backbone supply on the optimal locations of European hydrogen refueling stations?	Graph theory

10.1. Methodology

The theoretical framework brought graph theory forward as a method that fits the objective of this study. This was due to the properties of the theory, but also in part to the availability of a modeled hydrogen backbone for three decades. To employ graph theory, the Optimal Network Layout Tool will be employed. The methodology behind this tool will be explored and its use for this thesis will be highlighted. Thereafter, the assumed hydrogen backbone will be imported into the plain of the previously modeled stations.

10.1.1. Graph Theory

As its name indicates, graph theory deals with the study of graphs. A graph consists of a (finite) set of nodes and a (finite) set of edges, each connecting two nodes [119]. A graph can be used to represent any information that can be modeled in such a form and thus can be used for a wide variety of purposes. Generally, a node represents an object and the edge represents the relationship between these two nodes. In the case of this study, a node represents an HRS or a supply source of hydrogen. Hence, nodes can represent either sinks or sources. The edge would represent the physical connections between the HRS and the supply source, and thus be a pipeline. Figure 10.1 shows one of the simplest forms of graphs possible.

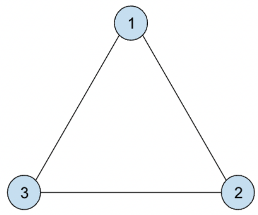


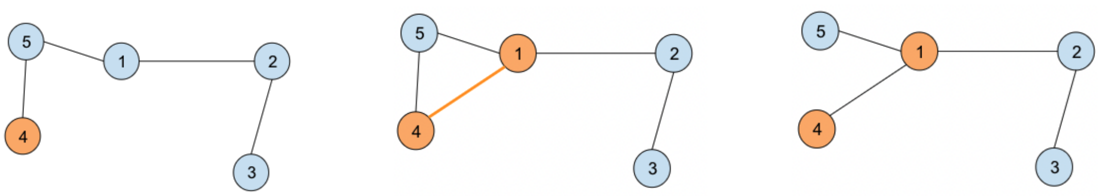
Figure 10.1: Simple Graph Example

One of the key features of emerging energy infrastructure networks is the desire to connect multiple sources and sinks in a cost-optimal manner, without redundancy [8]. This study shares this aim, as HRS are to be linked to a hydrogen backbone in cost-optimal manner. In 2030 and 2040, the assumed hydrogen backbone will consist of loose parts. Therefore, multiple source nodes are required for supply of hydrogen to the stations (sinks). Figure 10.3a displays an example of a set of demand and supply nodes. Cost optimality in such a network can be achieved by minimizing the weights of the edges, which would represent investment costs. To achieve such cost optimality, trees must be introduced. A graph is called a tree when no cycles are present [120]. No cycles are present, whenever there is exactly one path from each node to another. An undirected graph without cycles is called a forest. Hence, the goal of the network system design for 2030 and 2040 would be forests where 2050 would be a tree.

(a) Set of multiple source and sink nodes (b) Euclidean minimum spanning tree

Figure 10.2: K-Means clustering (random initialization) step 3a and 4

A category of trees are spanning trees, where all nodes are connected to one another. If a network consists of more than two nodes, multiple spanning trees exist. The minimum weight spanning tree is the tree among this set that has the lowest total edge weight, hence could be cost-optimal. There are many algorithms that aim to find minimum weight spanning trees [121]. Of these algorithms, that of Kruskal [122] will be used in this thesis as it is incorporated in the Optimal Network Layout Tool. To give an example, figure 10.2b shows a Euclidean minimum spanning tree. Such a tree equals the minimum cost spanning tree if the weights depend purely on their length. However, the edge weight for this thesis would depend on more pipeline aspects for hydrogen transport than only transportation distance. Therefore, the Optimal Network Layout Tool also makes use of a Delta change heuristic, which is incorporated due to its high performance and relatively low computational time [102]. One step of the heuristic is illustrated in the figure below.



(a) Minimum length spanning tree (b) Edge addition (c) Initial edge removal if costs are less

Figure 10.3: One iteration of the Delta change heuristic

The minimum length spanning tree, as just discussed, serves as the basis of the heuristic. In the following steps, the heuristic iteratively walks through two steps which are portrayed in figure 10.8c. Figure 10.3b shows how the heuristic adds an edge that connects node 4 to the closest node (1) that is not yet connected. This will iteratively happen for all nodes. The second step is then removing all other edges of the cycle that is created. If the removal leads to lower total costs of the network, the old edge replaces the new one. If not, the old edge will stand [8].

10.1.2. OPTIMAL NETWORK LAYOUT TOOL

The Optimal Network Layout Tool is a graph- theoretical tool developed at the TU Delft [123]. The tool is capable of optimizing energy network infrastructures that are multi-sink and/or multi-source while dealing with variable supply and demand patterns. The use of the tool is highly versatile since it is not intended for a specific case or geographic range. In addition, the ability to include existing connections allows this study to connect the HRS to an assumed hydrogen backbone. Huisman (2021) states that this makes the use of the tool preferable to other tools that are country-specific and cannot take into account existing connections. A comment that is equally relevant to this study.

For the tool to run properly, two well defined input sets are necessary: (1) a set of network nodes and their coordinates and (2) the supply and demand values of these nodes. Steps to define these are discussed in section 10.2.The Optimal Network Layout Tool works iteratively, starting with analysis of the demand:

Step 1: Analyse the demand - supply patterns
To represent fluctuations in supply and demand, the input of supply and demand values can be adjusted over different time steps. The three time steps for this study (2030, 2040, and 2050) should be treated seperately and therefore no supply and demand data for several time steps is present. This step is therefore irrelevant for this study.

Step 2: Determine representative set of k demand – supply profiles
The second step is to fixate the number of demand and supply profiles. Again, as only one timestep is analyzed each time, this visualisation is not relevant for this study. The one supply and demand set does however serve as the input for the generation of optimal networks.

Step 3: Determine minimal spanning tree
The network generated first is the minimal spanning tree. As previously discussed, this is the minimum-length network that connects all nodes based on Kruskal's algorithm. In the tool, just enough capacity is assigned to the edges in order to satisfy the demand asked by the nodes. The thickness of the edges can be used for visualisation of each respective capacity. The total costs are given as output in monetary units.

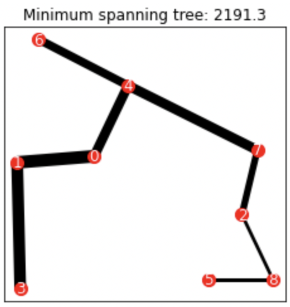


Figure 10.4: Minimal spanning tree

Step 4: Determine minimum-cost-spanning tree
As previously discussed, the minimal-spanning-tree solution does not take into account the capacity cost for building a certain connection. By rewiring the connections of the minimal-spanning-tree with the Delta change heuristic a better solution might be found. If no improvements are found, the minimum spanning tree is kept for Step 5. Figure 10.5 show the steps of profitable edge replacement on the example of 10.4.

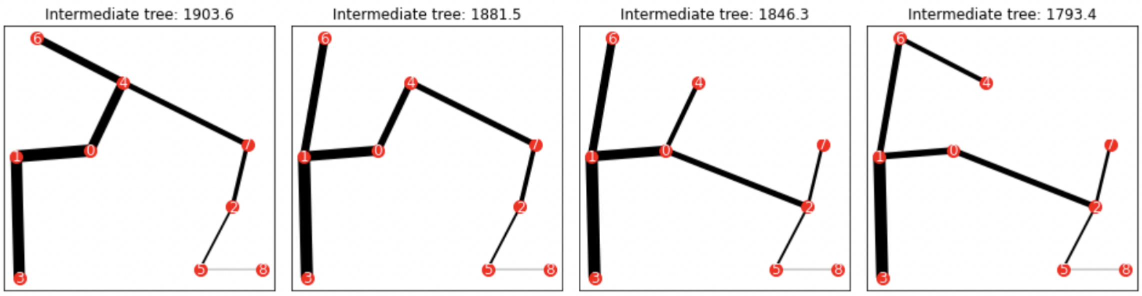


Figure 10.5: Intermediate steps of minimum-cost-spanning tree

Step 5: Determine minimum-cost-Steiner tree
By allowing extra splitting points on the edges, the network length can be shortened and the total costs of networks even further reduced. Such splitting points are called Steiner nodes and the resulting shortest length tree is called a minimum Steiner tree [102]. The Optimal Network Layout Tool searches for minimum-cost-Steiner tree, as not length but costs weight the edges. It must be noted that the addition of a splitting point in reality might also come with its costs. Therefore, a balance can be sought if costs per split are given. Figure 10.6 shows how Steiner nodes are added to the smallest angles in the tree one-by-one as long as cost improvements can be found.

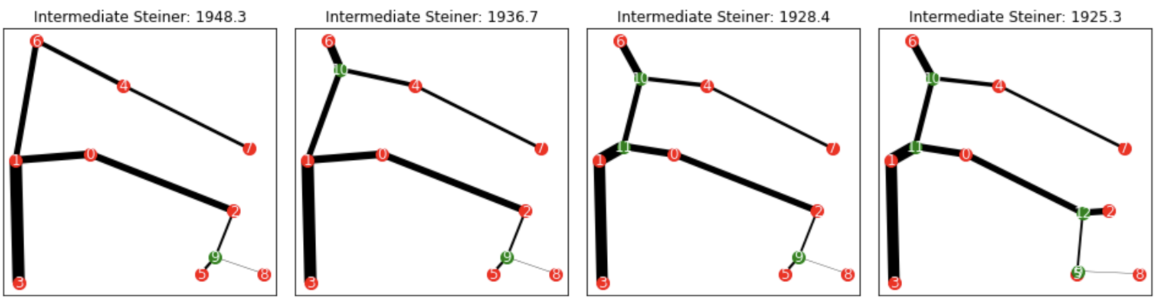


Figure 10.6: Intermediate steps of minimum-cost-Steiner tree

Step 6: Last improvement round
The final step of the model does not require much explanation as it is built for repitition of step 4 and 5.

RUN THE MODEL – WITH EXISTING CONNECTIONS

As previously introduced, the Optimal Network Layout Tool offers the possibility of modeling existing connections. These existing connections would function as to have "spare capacity". The same example as in the steps is displayed in figure 10.7, but now existing connections are present as blue edges.

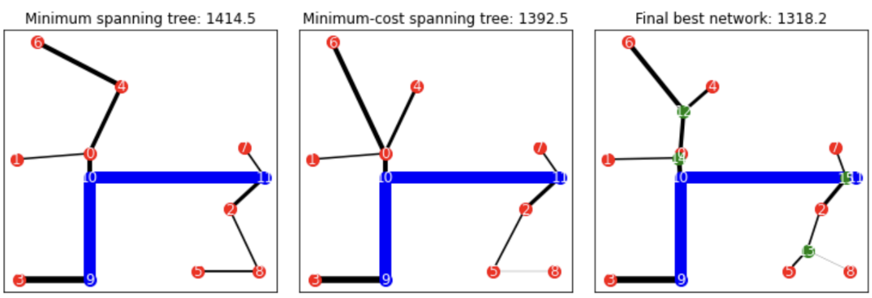


Figure 10.7: Network systems design with existing connection

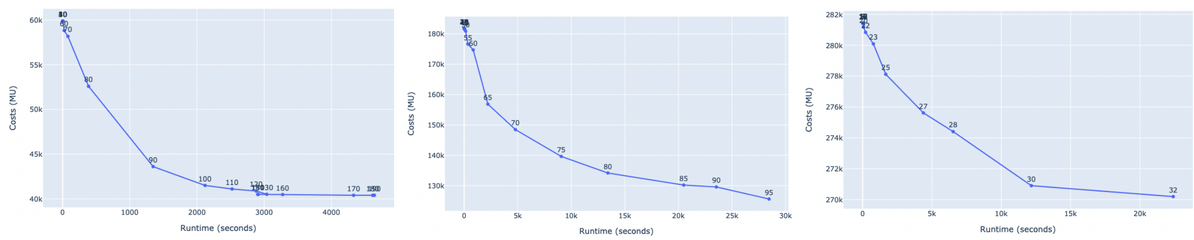
The figure shows that the same steps are executed to get the cost-optimal network topology. What is not the case in the example, but possible, is that the existing connection capacity can be expanded if it does not have sufficient spare capacity. Expansion costs are then given as output. Note that, if profitable, it is also possible that existing connections split. The costs of this split are modeled to be equal to the cost of splitting new connections.

Direct connection to existing connections
The existing connections are highly relevant for this study, as they can represent the presence of a hydrogen backbone. For modeling purposes, the HRS could tap into directly to this backbone. Therefore, it is important to state that, via Steiner nodes, the model also enables direct connection to a source via the existing connections (read backbone).

OPTIMAL NETWORK LAYOUT TOOL ADAPTION

The Optimal Network Layout Tool is a highly versatile tool, which can be used for a large range of purposes. However, for large networks some step might get computationally intensive. As this study deals with a very large network (European level of aggregation), problems with runtime are expected at step 4 and 5. Therefore, some small adaptions are made to cope with this.

Minimum-cost-spanning tree
The Delta change heuristic can become computationally heavy because a large number of possible edge changes are considered. Huisman aimed to reduce this by stopping the heuristic once the relative improvement fell below a certain threshold. However, due to nature of the heuristic, a small improvement in one iteration appeared to be no prediction for future improvements [8]. To cope with that, their research utilized a timer to prevent the heuristic from taking too long. A different approach is suggested for this study. The Delta change heuristic is iteratively looking for possible edge swaps. This translates into a search for a more cost-effective pipeline. However, for our geographical case, not every edge swap makes as much sense as the other. It would be impractical for the model to consider edge swaps over very large lengths: a swap over these distances is not cost-effective. Therefore, a threshold for the length of the edge to swap is modeled. It must be decided which threshold excludes mainly infavorable swaps. Therefore, the network costs (MU) are plotted against the runtime while the threshold is increased:



(a) 2030 (b) 2040 (c) 2050

Figure 10.8: Costs vs Runtime - minimum-cost-spanning tree

As can be seen, the thresholds differ per decade. The line indicate the threshold, so for (e.g.) 2030 a threshold for 90 km is set, which results in close to optimal costs and a runtime of 1200 seconds. The previous global optimum becomes a local one. However, results do not differ significantly.

Minimum-cost-Steiner tree
To reduce the runtime of step 5, the search for the minimum-cost-Steiner tree, no such method was found. Therefore, a timer approach is incorporated in the tool. If the tool takes too long to find an improvement, the step gets stopped. Although a timer may reduce the performance of the heuristic, it allows control over the computation time.

10.2. DATA PROCESSING

As introduced the tool needs two datasets to run: a set of network nodes and their coordinates and the supply and demand values of these nodes. These datasets need to be structured in the right format for the tool to process. Therefore, all data gathered (HRS locations and the backbone) must be processed properly to create an optimal network design. Note that, for the tool to work, the amount of demand and supply must be equal at all time. A more technical explanation of how demand and supply are equaled is given in Appendix F. The other forms of data processing are explained in the following sections.

10.2.1. SUPPLY NODES AND EXISTING CONNECTIONS: THE HYDROGEN BACKBONE

As previously identified, the incorporation of an assumed hydrogen backbone in an optimal network system design should be done in the form of existing connections. The exisiting connections should function as a transportation medium, fed by a source (node) of hydrogen. These source nodes should be placed such that direct connection from a demand node is not possible. To ensure this, the supply nodes are placed exactly on the end of each (piece of the) backbone. As such, connection to the source of hydrogen will always take place through the backbone.

Capacity of the backbone

Since the backbone will be present to facilitate a large group of hydrogen purposes, it is assumed that the capacity will always be sufficient to serve heavy-duty trucks. Therefore, the backbone is modeled such that the "spare capacity" is always large enough to meet demand of the nodes attached to the existing connections. Hence, the capacity of the backbone is modeled that it's large enough to capacitate the sum of all demand nodes.

Adapting Huisman's backbone

The availability of the backbone is not instant, as it must be imported and adapted to the results of the previous chapters to be of use. The NUTS regions that Huisman used as centers for the backbone have been adapted to the purpose of his thesis. The "growing" backbone in figure 10.9 is the result.

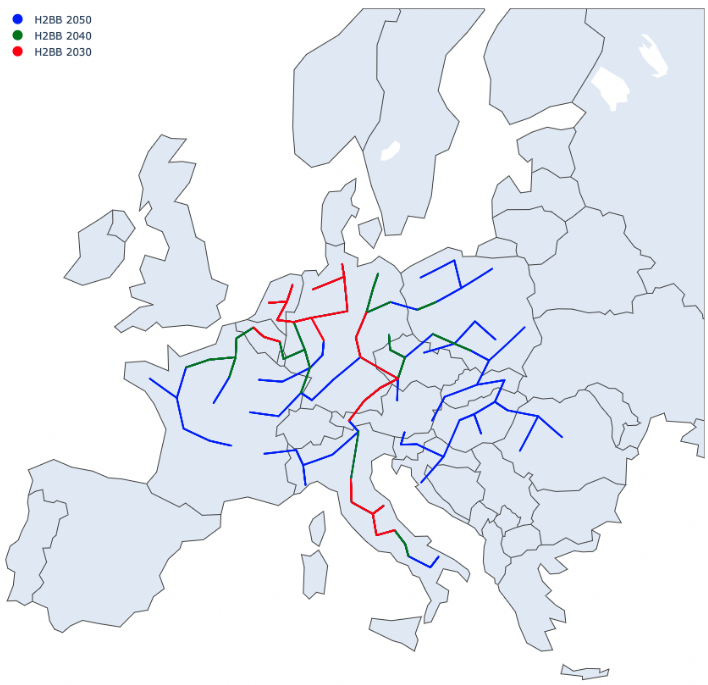


Figure 10.9: Hydrogen Backbone: 2030, 2040, 2050

10.2.2. DEMAND NODES: HRS

To enable use of the Optimal Network Layout Tool, the HRS locations and the sizes that result from the station location modeling should be incorporated as demand nodes.

In reality, supply of hydrogen by backbone does not make sense for all stations. For modeling efforts, it is assumed that all locations are eligible for connection to the backbone and that investment efforts are equal all over Europe. However, to get closer to reality, only stations within a certain distance threshold are included in the modeling. This distance threshold is already assumed to be 100 km in section B.3.2 as only distribution pipelines are considered. Stations that lie further than 100 km of the backbone (euclidean, "as the crow flies") are left out of range and will thus not be included in further modeling. Figures 10.10 and 10.11 display which stations fall in this range for 2030 and 2050 respectively.

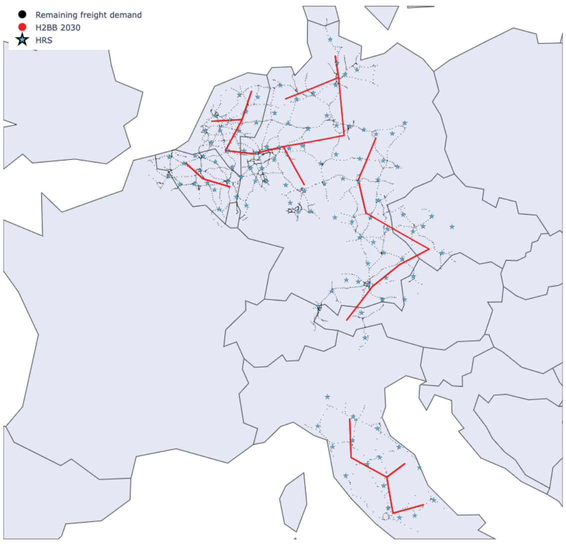


Figure 10.10: HRS network in range of 2030 backbone

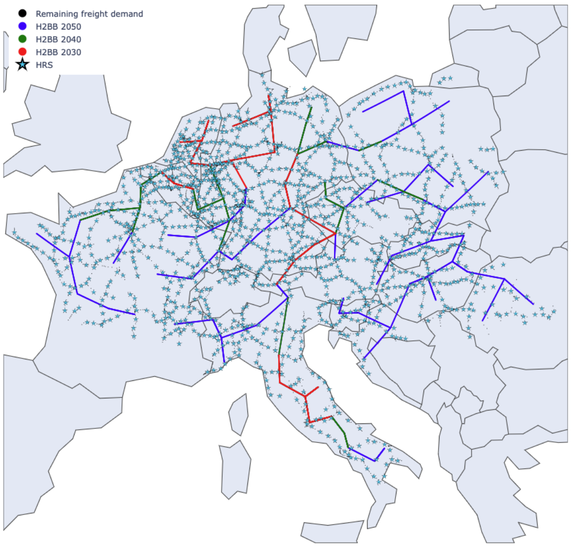
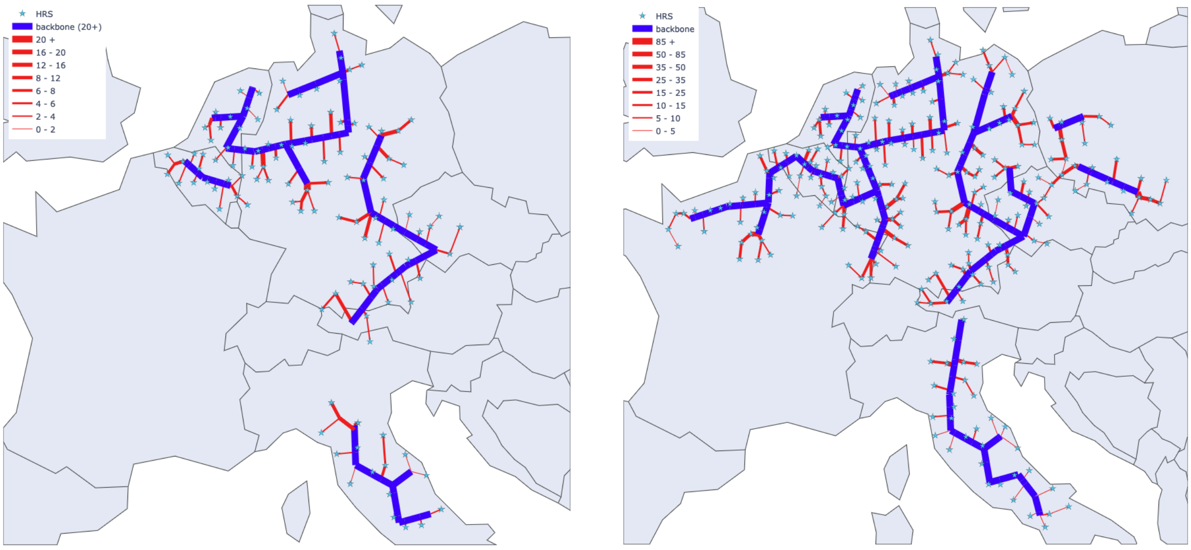


Figure 10.11: HRS network in range of 2050 backbone

Note that this leaves only these models for inclusion in the network system design and thus eligible for further optimization in the final step.

10.3. Initial application

Initial application of the tool yields the results for 2030, 2040 and 2050 presented in figures 10.12b and 10.13. As indicated in the legend, the blue line represents the backbone and the red lines the distribution piping.



(a) Initial network 2030

(b) Initial network 2040

Figure 10.12: Initial network system design

The figures 10.12a and 10.12b display that each backbone is able to supply the connected refueling stations. Otherwise, the model would seek connection between the different parts of the backbone. Figure 10.13 displays the 2050 and most complete network.

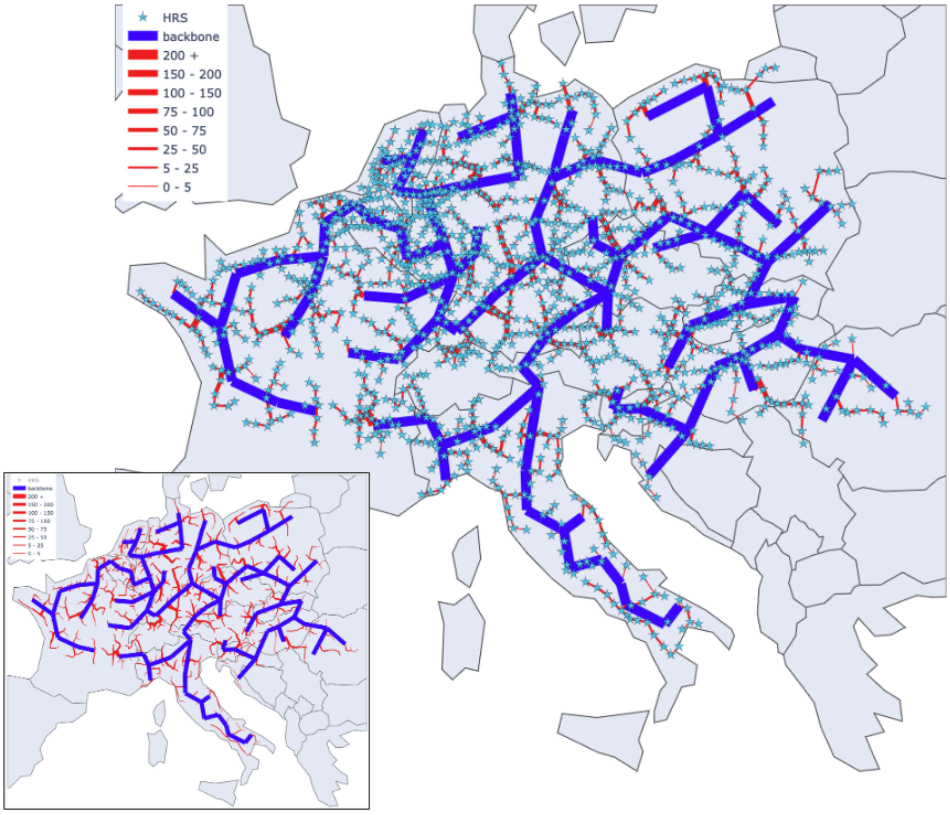


Figure 10.13: Demand-based optimal network design 2050 and smaller visualisation of the network without HRS

A first look at the results shows a network that is growing over time. Moreover, it is notable that by 2050 a relatively larger network is present. In chapter 12, these results will be further discussed and analyzed.

10.4. CONCLUDING REMARKS

This section is based on the output of the station location model (discussed in chapter 9). The aim of this chapter is to connect these demand-based station locations to a supply source. This supply source is modeled as a backbone and the connection to it is modeled as distribution pipes. By using an Optimal Network Layout Tool, an optimal design for this supply network is created. This leads to the fact that there are now demand-optimal station locations (a topology) and a supply-optimal network of pipelines connected to those stations. This system is made for three decades and is called the Network System Design.

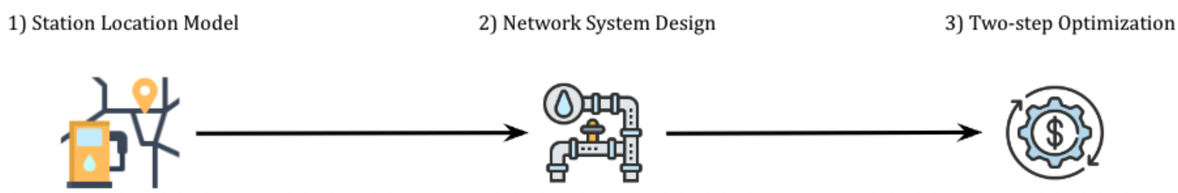


Figure 10.14: Skeletal structure theoretical framework

The next step in the two-step optimization process is called "the two-step optimization". To further optimize the topology of refueling stations, the costs of supply are taken into account in station location determination. The topology should be influenced by the costs of piping. This leads to further optimization in terms of CAPEX, as indicated in the performance measures.

11

Two-step Optimization

This section presents the methodology and initial results of the two-step optimization. Section 11.1 lays out the methodology to determine the optimal location of an HRS based on demand and supply of the previous chapters. The sub-question that this section supports to answer is identified in the table below. Optimizing the locations that were based on demand with supply will show the influence of a potential backbone. Initial results will be presented in section 11.2 and will be further analyzed in chapter 12.

Table 11.1: Research Goals of Network System Design

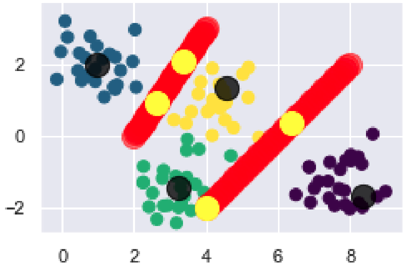
Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion addressed in this section	**Research Method**
What is the influence of hydrogen backbone supply on the optimal locations of European hydrogen refueling stations?	Two-step Optimization

11.1. Methodology

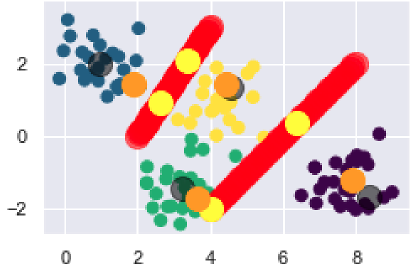
The two-step optimization will make use of the models and methodology that have been previously discussed. The aim is to search for a balance between the output of the station location model and the network system design. Previously, the station locations were only influenced by demand. The location determination was based on a weighted average of datapoints of demand, which were clustered to get covered by European network. Thereafter, a network system design was created and each HRS (within range) was connected to supply of a backbone. So far, this supply has had no influence on the location of the HRS. Hence, the station location is sub-optimal and only optimized towards demand. To search for the influence of supply on the station locations, the distance towards the hydrogen backbone is considered. This distance is opposed to the distance of the HRS to all demand datapoints in the cluster of the HRS.

11.1.1. Average weighted location

To determine a new station location based on the balance of demand and supply, a weighted average is used. Previously, the k-means algorithm determined the station location based on the weighted average of demand points. In this step, each cluster per HRS is isolated and the point to supply of that HRS is added to form a new set. Based on the weighted average of this new set, which thus consists of a demand points and a point that connects to a supply source, a location is determined. Figure 11.1 shows an implementation of this methodology based on the same example as before. The data points in each cluster are given a weight of 1, except for the previously identified points which have been given a weight of 20. These were given this weight to show the influence of the weights in the location determination. The supply point (yellow dot) on the backbone, is given a weight of 30. It can be seen that the new station location (the orange dot) moves linearly toward the backbone. Since the yellow and purple clusters had larger weights of the data points, the new location is less affected.



(a) Determine closest point of supply



(b) Location based on weighted average

Figure 11.1: Simple example to determine location based on demand and supply

Applied on the dataset of this study, the weights of the datapoints represent the demand for hydrogen per demand node. The supply points are either points on the backbone, or other stations which subsequently connect the HRS to the backbone.

11.1.2. WEIGHTS FOR BACKBONE AND DATAPOINTS

For the initial station determination, the weights of the datapoints were the vehicle kilometers of FC-HDTs which could be translated to a demand for hydrogen. The same values no longer apply to the weighting of this second optimization. In order to arrive at the correct weights, it is necessary to be clear what the costs are of making the HRS move in a particular direction. These costs are composed of demand and supply costs and simplified to relative distances. A balance can be found in the costs per year per kilometer.

SUPPLY COSTS

First, the cost of supply is determined, since it is less complex than that of demand. There is only point of supply that is added to the set of nodes. The distance of the HRS to this supply node make up the relative costs. As determined in the system description, 1 kilometer of semi-retrofitted pipeline has a CAPEX of 0.48M€ / km. For the sake of simplicity, the OPEX has been excluded. The depreciation period for this kilometer of pipeline has been set at 30 to 55 years. For modeling purposes, this is scoped to 50 years. As such, the costs per year per kilometer to move away (or towards) the supply point is 9600€ (480000 / 50). This value can be used as a weight for each connection point to a supply source.

DEMAND COSTS

The cost of demand is somewhat more complex. To be comparable with the costs of supply, the costs per year per km of movement from the HRS should be calculated. In practice, the relocation of an HRS would result in trucks (that refuel there) having to drive further or closer if their tank is empty. Any kilometer these FC-HDTs would have to drive extra, would cost them extra fuel. Other aspects, such as an increase in the drivers' wages, are negligible and thus not considered. Hence, the price to fuel a FC-HDT for 1 km determines the cost per demand node. The system description revealed that a FC-HDT can drive 17 km on 1 kg of hydrogen. It was further assumed that 1 kg of hydrogen will cost 4€ per kg in 2030, 2040 and 2050 (due to subsidies and the necessary diesel parity). This leads to the assumption, that it costs 4/17 € for every kilometer a FC-HDT has to drive extra.

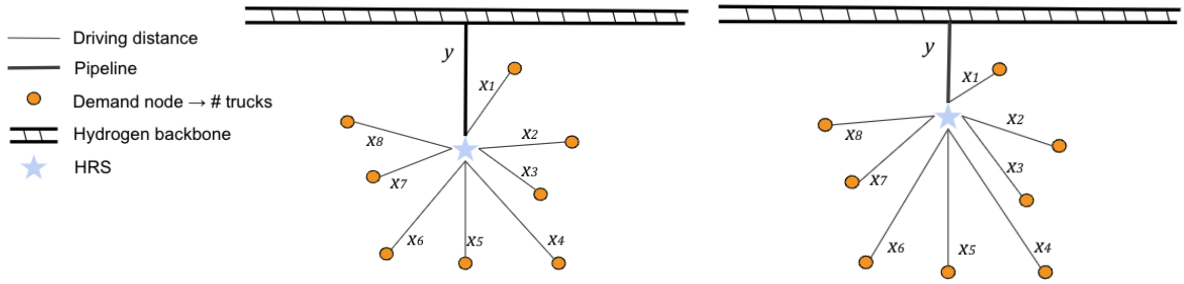
This cost per kilometer, per truck, per year represents the weight from demand towards the movement of an HRS. However, since the data points were created by recalculating a flow to points, no clear number of trucks per data point is known. Previously, the datapoints were appointed values that represent a demand for hydrogen in kilograms per year (calculated based on vehicle kilometers). Put in other words: the amount of hydrogen that is "used" on that part of the highway over 1 year. This demand for hydrogen could be therefore be translated to a representative of a number of trucks, in need to refuel at that datapoint over a year. This can be done by dividing the hydrogen demand per data point with the value a full tank of a FC-HDT. This simplification may feel counterintuitive, but as a network the assumptions hold. As the fuel storage of hydrogen is limited to 85 kg per truck (700 bar), the demand for hydrogen per node must be divided by 85 to get a number that represent a number of truck refuelings per year.

BALANCING THE COSTS

All by all, the aim of balancing the weights of demand and supply can be defined by a mathematical formula with the objective to minimize:

$$\min\left(W_t \sum_{i\in C_k} d(i, c_k) + W_p d(c_k, b_k)\right)$$

The minimization takes place over the addition of the weighted distances of supply and demand (d). Ck represents the cluster from centroid ck, and bk stands for the connection point of the backbone. The weight of supply is given as Wp and the weight of demand by Wt. Figure 11.2 visualizes how this process should be interpreted.



(a) Determine closest point of supply

(b) Location based on weighted average

Figure 11.2: Simple example to determine location based on demand and supply

11.1.3. OPTIMIZATION PROBLEM: INTERDEPENDENCY

The example in figure shows possible movement of an HRS connected to the backbone due to distance of supply. This type of optimization is perfectly possible for standalone HRSs directly connected to a supply source. However, when a network of HRS is analyzed, mathematical complications come into play. Not every HRS is directly connected to the backbone, but many are interconnected to reach the supply through another station. This causes interdependency problems in optimization, because if one station moves, the supply source of another moves. On top of this comes the problem that stations may move in such a way that their cluster is no longer representative. This type of optimization is too hard to compute in the current setup.

11.2. INITIAL APPLICATION

Since the interconnected pipelines of the network system precludes the use of two-step optimization, the model is applied to a different network: the HRS locations, as determined by demand, "directly connected" to the backbone. This means that no station is modeled as "interconnected," so the supply network is not optimal. While this design is not an optimal network, it still allows for recommendations on the influence of supply on station location decisions. Especially since many stations are in close proximity to the backbone (< 100 km) and thus directly connected.

The exact method of section 11.1.1 is applied. By taking the weighted average of the defined weights of the demand points and the backbone connection we get figures 11.3, 11.4, and 11.5. The blue stars on the maps represent the "old" HRS locations, as determined by the station location model. The orange stars represent the "new" locations. These locations are based on the balance between costs of supply and demand as visualized in figure 11.2. The red line displays the backbone.

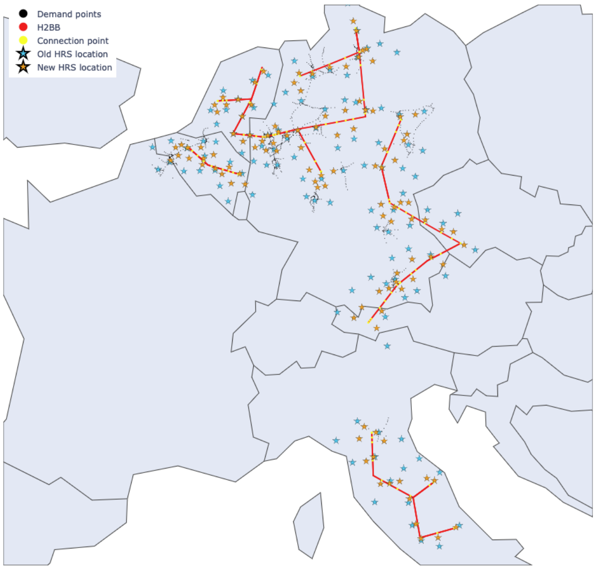


Figure 11.3: Two-step optimization 2030

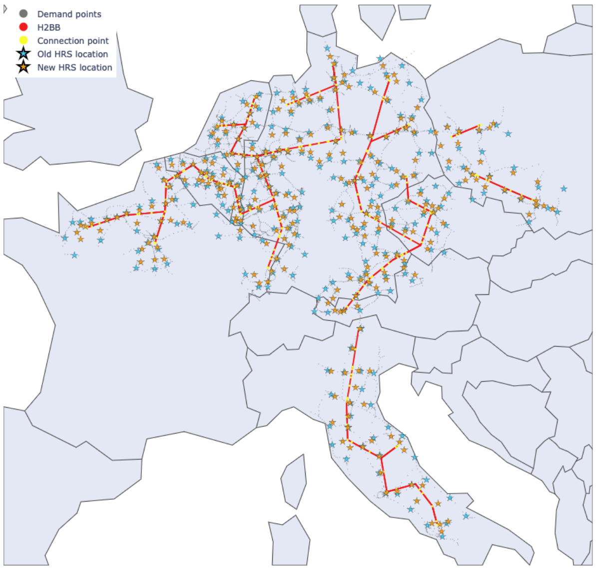


Figure 11.4: Two-step optimization 2040

An initial inspection of the figures makes clear that the cost and supply figures appear to be relatively well balanced. The HRS do move, but their new locations are not exactly on top of the hydrogen backbone. As is inherent to the model, the stations move in a straight line towards the backbone

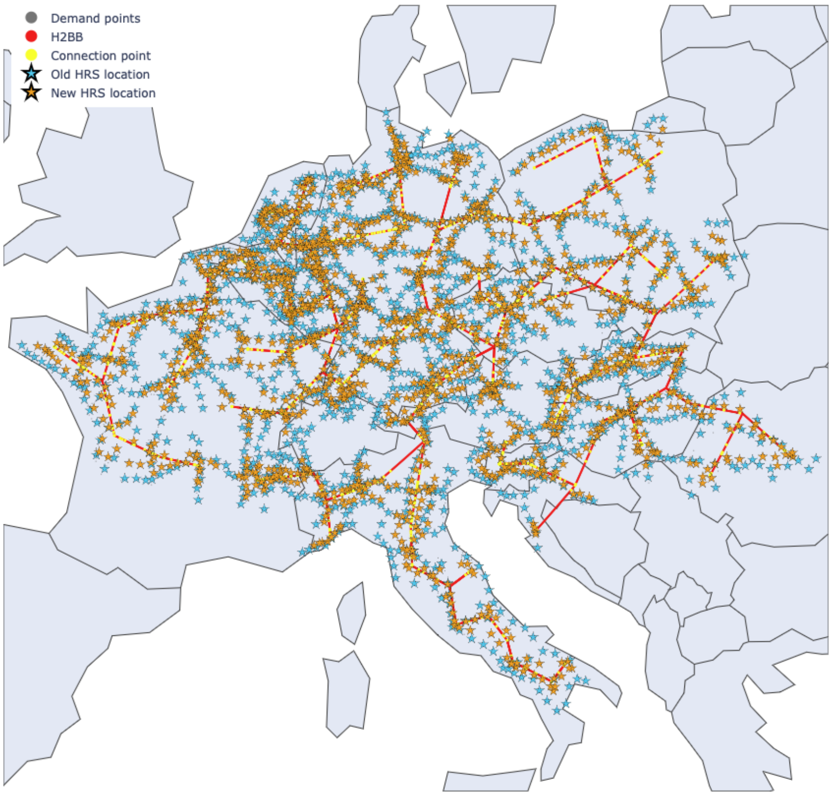


Figure 11.5: Two-step optimization 2050

The two-step optimization seems to do what is asked: the demand based topology is still intact, but the influence of supply is clearly visible. Chapter 12 will dive deeper into these results.

11.3. CONCLUDING REMARKS

This section, the two-step optimization, presents the final step of the modeling process. The network system design is supposed to function as input, as displayed in figure 11.6. The network system design has an optimal line of supply for each demand-optimal station location. However, the locations of the stations are not yet influenced by supply. This final step aims to do so by relocating the stations such that the costs of demand and supply are balanced.

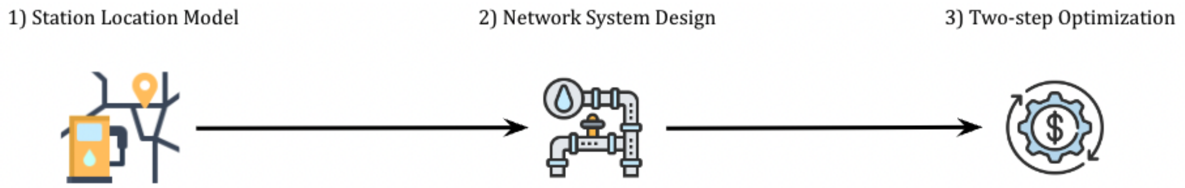


Figure 11.6: Skeletal structure theoretical framework

Unfortunately, it is infeasible for this study to consider the optimal line of supply of the network system design in the two-step optimization. This is due to an optimization problem inflicted by interconnection. For that reason, all stations are assumed to be directly connected to the hydrogen backbone instead of being part of an interconnected network. Each station thus has its own pipe towards the hydrogen backbone for supply. This network is optimized for 2030, 2040 and 2050. This results in three maps that show demand and supply optimal locations in their own fashion. When interpreted correctly, the results presented still have a lot of value.

The following chapter analyzes the results that have been built up over the past four chapters. This is followed by an analysis of scenarios and uncertainties prior to conclusions, discussion and recommendations.

12

Analysis of Results

In this chapter, the methodologies discussed will be fully applied and the results analyzed. The analysis will be performed chapter by chapter in the same order as already presented. First, the freight demand model will be discussed and the regulatory basis will be analyzed. Thereafter, the optimal network of HRS is discussed by station location model, network system design and two-step optimization. To make analysis possible, the four performance measures are consulted again. In this way, clear analyses can be conducted to formulate recommendations and reach final conclusions.

Following the analysis, key assumptions and uncertainties underlying the modeling steps are discussed in section 12.2. In the final section, 12.3, an attempt of validation is provided. In the next chapter, the sensitivity of the inputs resulting from the system description will be tested using scenarios.

12.1. Analyses

In this section, the results of each modelling step will be touched upon and analyzed. A general explanation of the results will be discussed and notable discrepancies will be highlighted. As a reminder, the four performance measures to assess a network of refueling stations are repeated:

- Utilization rate of HRS
- Weighted average distance (travel time) to HRS
- Maximize total vehicle kilometers covered
- Minimize CAPEX (number of stations)

These criteria are used to set up the methodology and are therefore already present in the output of the models. They will be used to guide the discussion and key findings.

12.1.1. Freight demand modeling

The results of the freight demand model gave a clear overview of where trucks were located over a year. The traffic flow, translated to datapoints, revealed initial insights of a busy center of Europe with peripheral areas with less heavy-duty traffic. The freight demand model was intended primarily as a basis for further analysis, and does not itself require analysis. However, the results are validated to confirm that this basis of modeling coincides with real truck traffic counts in section 12.3.

12.1.2. Analysis of Regulatory Basis

As discussed in chapter 8, the AFIR is analyzed on two out of the four performance measures: (1) potential under- or overutilization of the mandatory station locations and (2) their coverage of demand in Europe. For both analyses, the AFIR locations and the resulting freight model network are placed in the same model plane.

AFIR analysis: potential under- or overutilization

The demand for hydrogen from the freight demand model is allocated to the stations located by AFIR as described in chapter 8. No demand nodes (datapoints representing trucks) outside of 150 km range of a station (e.g. UK or Ukrain) are in scope, so it is reasonable to assume that the trucks would refuel at the stations closest. The range of a FC-HDT is assumed to be 800+ km. The allocated demand to each station displays potential under- or overutilization as it reveals how much demand the 2 tonnes per day stations could potentially capture. To highlight under- and overutilization, the HRS in figure 14.2 are displayed as nodes with coloring. If the node is underutilized (0 - 0.5 tonne/day), it is colored red, which gradually changes to orange, yellow and green, depending on the potential demand. Anything larger than 2 tonnes per day is colored darker green as overutilization would take place and more refueling capacity is necessary. It is still colored green as overutilization is not necessarily a problem: HRS capacity may be scaled up, but the minimum is determined by AFIR and thus cannot be lowered.

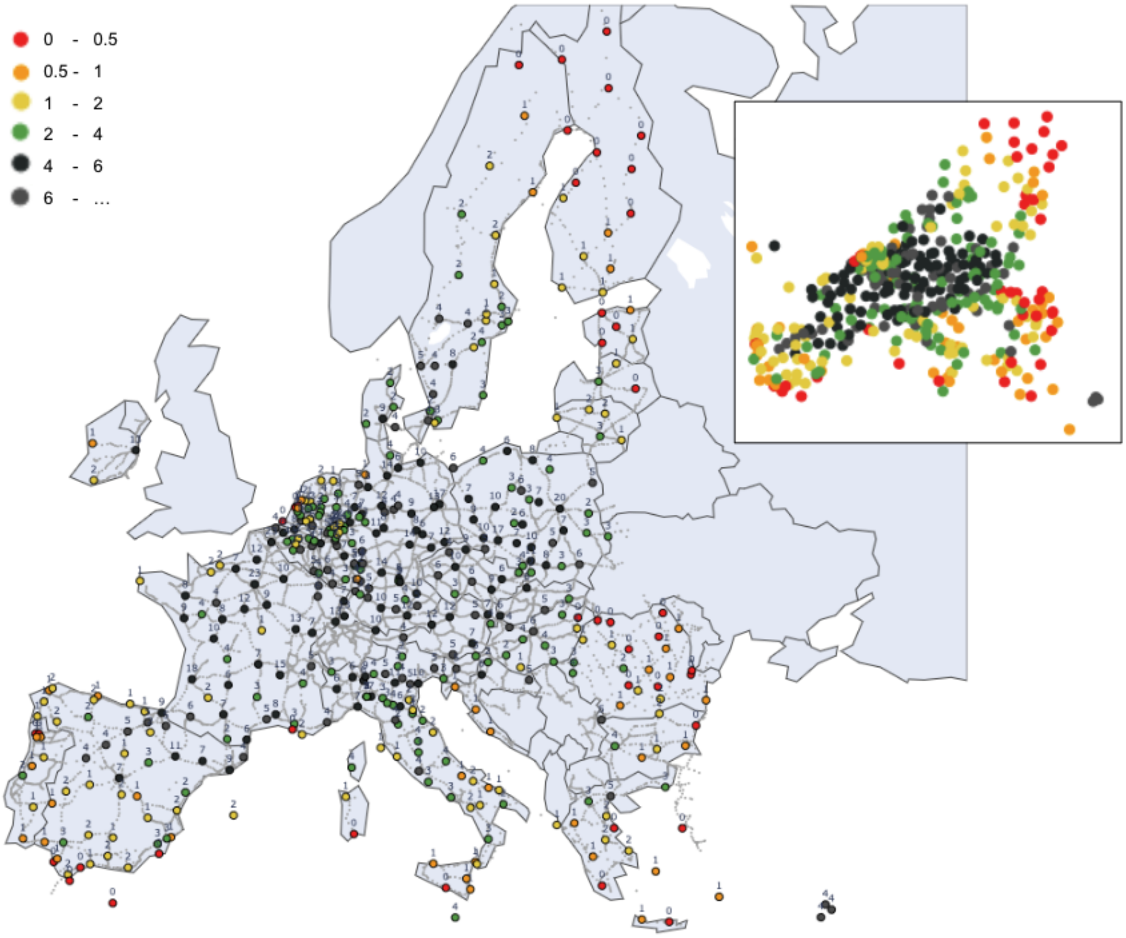


Figure 12.1: Demand-based network - indicative coloring for station size

Figure 14.2 shows that central Europe will easily acquire enough market traction to cover the two tonnes of hydrogen per day per station. This will not be the case in the more peripheral areas because there are far fewer trucks there. It should be noted that only heavy-duty trucks have been taken into account, while light-duty vehicles may also play a role. For the purpose of this thesis, this has been disregarded, but it is an important fact for the discussion of the policy proposal.

AFIR analysis: coverage of demand

Besides analyzing the utilization of the station locations, analysis of the capacity prescribed by the EU is also useful. Chapter 8 prescribes use of a p-median model to analyze the coverage. The model searches for the closest demand points in its proximity and adds them to a list of "served" demand points until the stations capacity is reached. The results of this analysis are displayed in figures 12.6 and 12.7.

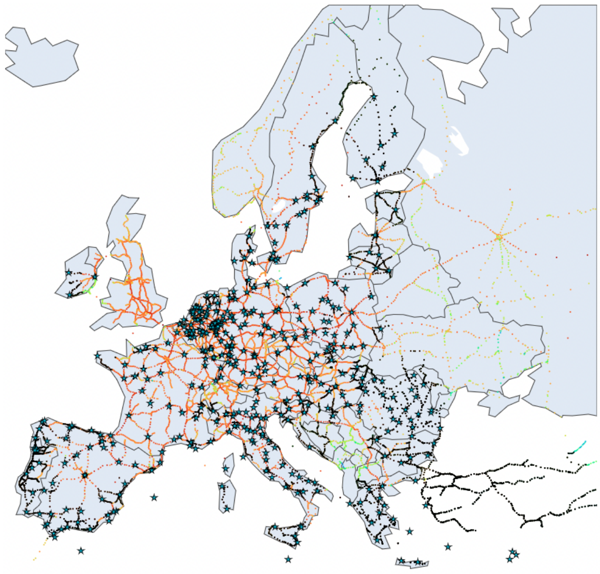


Figure 12.2: AFIR network and covered demand (black) 2030

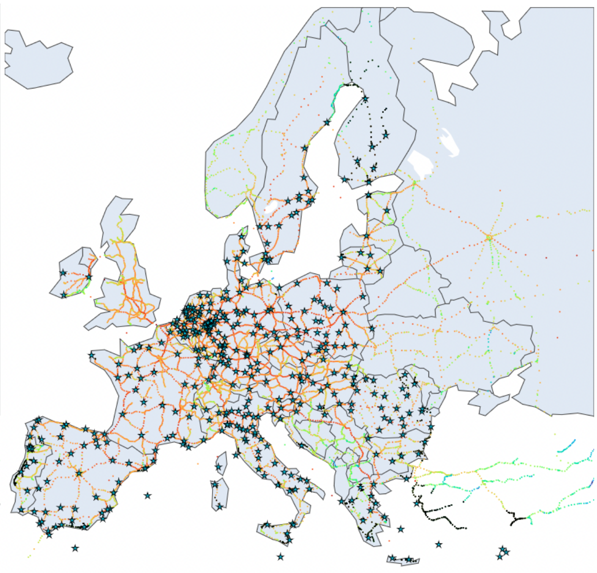


Figure 12.3: AFIR network and covered demand (black) 2050

As the figures show, the initial AFIR network already serves a fair portion of the network. Calculations reveal that 44% of total demand in 2030 is covered by the 2 tonne per day stations. A comparison of the demand covered in 2030 and 2050 shows the increase in market share of FC-HDTs. The larger the number of vehicle kilometers, the less datapoints (or trucks) the stations can serve under strict capacity. Calculations even show the total coverage drops from 44% in 2030 to 10.9% in 2050. However, as the AFIR is written for 2030, the analysis for 2040 and 2050 are of less value.

KEY FINDINGS AFIR ANALYSIS

The two analyses of AFIR reveal a number of interesting results that can be taken into account for policy recommendations. These are presented in an itemized list for the sake of overview:

- *The addition of HRS every 150 km, on top of the proposed urban nodes seems unnecessary.*
  In section 9, the methodology already revealed that very few HRS have to be added to fulfill the requirement of having an HRS present every 150 km along TenT corridors. The analysis that followed even revealed that many of these stations (e.g. Norway) do not capture much demand and are therefore not essential for the network.
- *The HRS in peripheral areas of Europe are headed for underutilization.*
  The allocation of demand towards the AFIR locations clearly shows a pattern. Figure 14.2 displays how the peripheral areas in specific will become underutilized.
- *Stations in central Europe do not nearly cover all demand at 2 tonnes per day.*
  Figure 14.2 already signalled overutilization of most stations in central Europe. On top of that, figure 12.6 shows how little of the central Europe highway network is captured by 2 tonnes per day stations. However, it must be noted that 44% of overall demand does get captured in 2030.

12.1.3. STATION LOCATION MODELING

The application of the station location model resulted in the topology of demand-optimal station locations presented in chapter 9. This topology, based on weighted k-means modeling, is used as basis for the analysis. The station locations each have a cluster of datapoints assigned to them, representing a demand for hdyrogen. To get insights into the market, the hydrogen demand that each cluster possesses is summed and added to their respective HRS. In this way, three analyses can be conducted that allow for recommendations; (1) analysis of the distribution of stations over a set of given station archetypes (stations of standard sizes); (2) analysis of the necessary number of stations and locations to reach a critical point of demand coverage; (3) visual analysis of the demand allocation over the network to get insights into the market.

As discussed in chapter 5, the demand is of critical importance in the initial phase of the network. This became apparent by the ranking of modeling abilities by Shell in Appendix E. Shell ranked the distance for trucks above all other measures because they want to gain early market traction. Therefore, for the demand-oriented station location model, only the early phase 2030 results are analyzed.

Analysis: spatial coverage of stations

By definition, the weighted k-means algorithm resulted in a minimal (weighted) distance from each demand point to an HRS and 100% coverage of the vehicle kilometers. It therefore optimizes the demand performance measures (coverage and weighted distance). To quantify performance of the algorithm, the distance from each demand node to an HRS is calculated. These results are presented in figure 12.4.

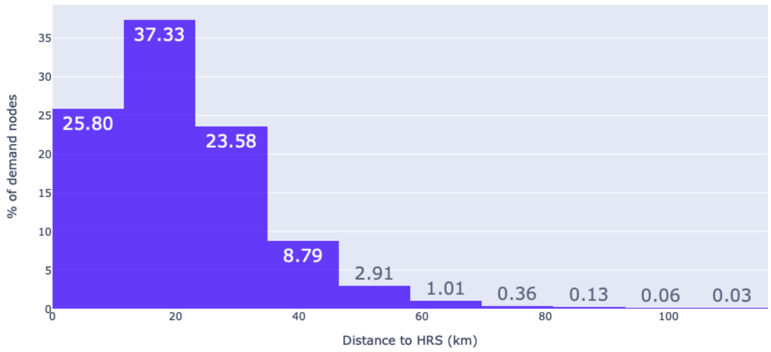


Figure 12.4: Distribution of distance from demand node to HRS

The figure shows that the network fully covers the demand in good fashion: more than 99% of all nodes are served within 70 km Euclidean distance. This network is presented in figure 12.6. The intention of the station location model was to mainly account for demand. However, as introduced in the methodology in section 9, the performance measures of supply (CAPEX and utilization) can also be reduced in this step. By excluding stations from the network, an initial balance can be sought. An optimal network does not have to cover 100% of demand if CAPEX can be greatly reduced by building less stations. Moreover, stations with a low utilization rate can be excluded further optimizing the network in terms of the performance measures. This balance is sought by iteratively lowering the threshold of when an HRS is omitted and measuring coverage. Starting at a minimum demand of 6 tonnes per day, ending at 0 tonne per day (no threshold). Figure 12.5 shows the results

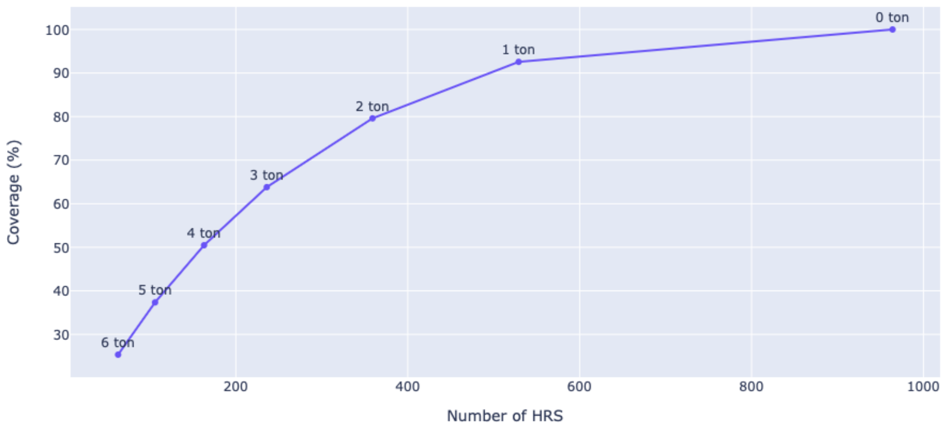


Figure 12.5: Threshold of minimum station size for coverage

As it should be, with no threshold, the graph shows full coverage (100%) with 964 stations. The largest step in efficiency is made by excluding stations with potential demand lower than 1 tonnes per day. The number of stations get reduced by almost half (529 HRS), while losing less than 10% of coverage (92.5% coverage). The threshold larger than 1 tonnes per day are less efficient and thus not considered. The topology of hydrogen refueling staions with more than 1 tonne of expected demand is displayed in figure 12.7

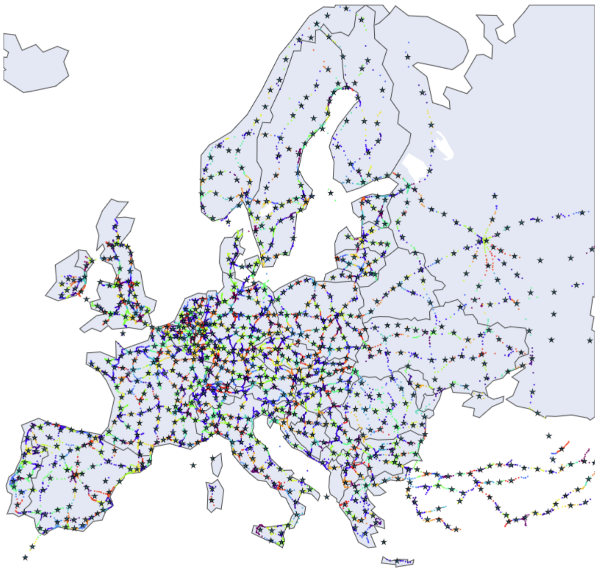


Figure 12.6: Optimally distributed network of HRS 2030

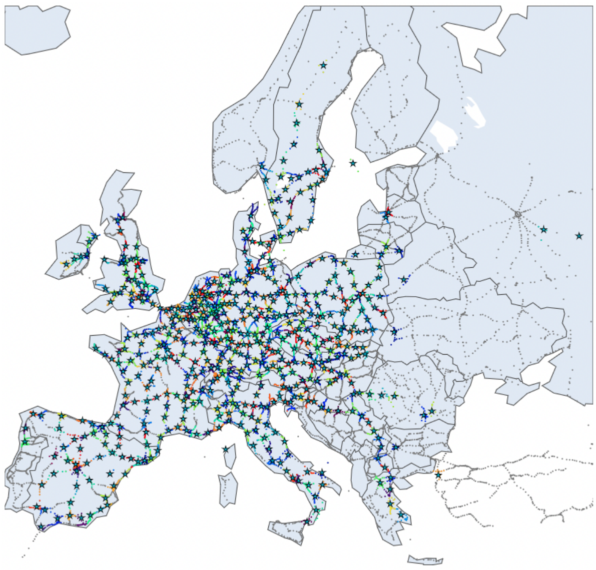


Figure 12.7: HRS network of stations > 1 tonnes / day

Since this network loses coverage only slightly and the CAPEX cost falls sharply, it may be a better basis for further modeling in the two-step optimization. First, however, the other performance measure of supply can be tested by analyzing the different archetypes and their utilization.

Analysis: archetype distribution

As weighted k-means is used to distribute the stations, the total number of stations is pre-defined. The station location modeling section (9) stipulates that this number should be based on an average station size and total demand. This resulted in 964 stations with an average size of 2 tonnes per day for 2030. Each specific station size however, is not yet determined. Only the expected demand per station is calculated. Therefore, three archetypes of stations are drawn up in consultation with Shell. Previously, the reasons for predicting 2 and 6 tonnes per day stations have been touched upon. A 10 tonnes per day station is added to this list as it is expected to add value in future cost reduction. In figure 12.8 stations are added to bins where demand grows larger with 1 tonne per bin (0 to 10 tonnes / day). Coloring each bin up untill the three archetype capacities, results in the predicted amount of stations per archetype.

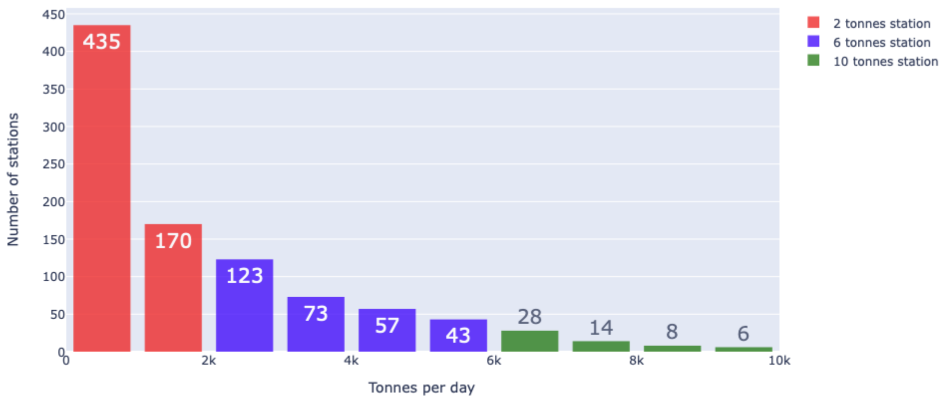


Figure 12.8: Number of stations per archetype: binned histogram

Using this method, utilization rates for these archetypes are consequently predicted: the bins show what part of the archetype serves serve how much demand. The 6 and 10 t/d stations enjoy rather good utilization rates as they by definition have a minimum of 2 and 6 t/d of demand respectively. The modeled 2 t/d stations show very low utilization rates, as 71.9% utilizes less than half of their capacity (1 t/d). Eliminating these 1 tonne per day stations from the network would cause a large increase of the overall utilization rates.

This finding appears to be in line with the result that stemmed from the CAPEX reduction of building less stations. Exclusion of the stations with less than 1 tonne per day of expected demand leads to such a gain in efficiency on the supply side, that this network is assumed the basis for the remainder of this study.

Analysis: Visualisation of demand allocation
As a final part of the analysis of the demand-determined network, a new visualization is presented. Figure 12.9 shows the network of stations greater than 1 tonne per day and colors them according to their potential demand. Representing the network in this way will highlight regions of interest to Shell and help identify key corridors. These insights can act as a solid foundation for network planning and help prioritize the construction of hydrogen refueling stations.

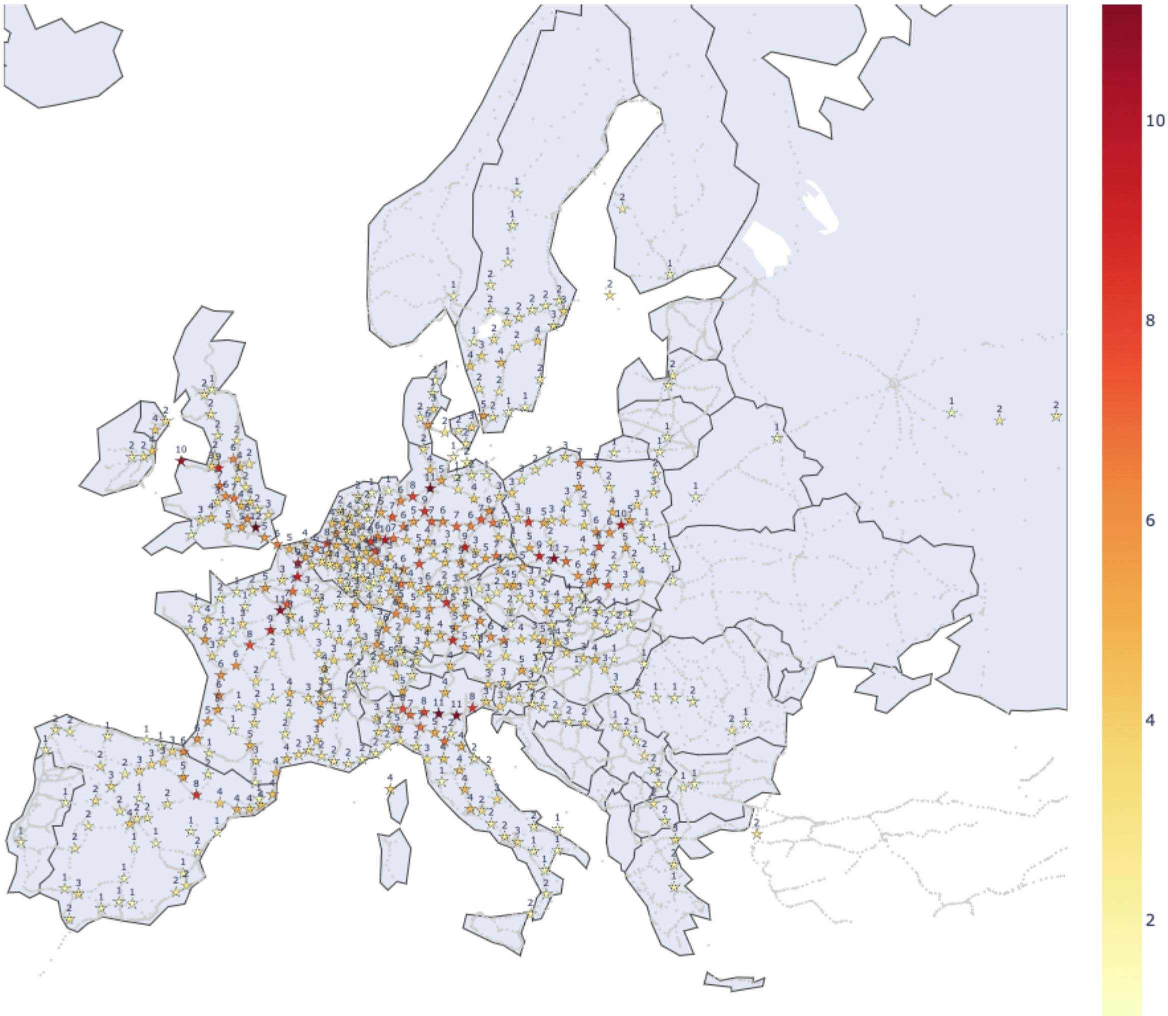


Figure 12.9: AFIR network - allocated demand to show over- and underutilization

Based on visual inspection, this figure already enables some high level takeouts:

- Highest overall potential in DACH (Germany/Austria/Switzerland), Benelux and UK.
- High potential on key corridors in France, Italy and Poland.
- Relatively low potential in Spain, Denmark and Sweden.
- Norway, Portugal and Finland should be considered countries of non-interest.
- Interesting key corridor through Eastern Europe to Greece.

For a more in-depth analysis at the (e.g.) national level, it is recommended these results are exported to external network planning tools. These can (and will) be used for network planning by Shell.

KEY FINDINGS STATION LOCATION MODEL

The three analyses of the demand based network reveal a number of interesting results that can be taken into account for recommendations to Shell. These are presented in an itemzied list for the sake of overview:

- *The three archetypes of 2, 6 and 10 t/d are predicted to cover a realistic network of HRS*
  For 2030, the three proposed archetypes cover the whole range of predicted HRS sizes. However, underutilization is expected to be a problem for the smaller stations.
- *Excluding stations that are expected to recieve less than 1 t/d of demand leads to a large increase of utilization rates while coverage does not drop much*
  The exclusion of these stations will leads to disposing 435 stations with a utilization rate lower than 50%. This would exclude 71.9% of all 2t/d HRS. The network of HRS would then no longer ensure full coverage, however 92.5% of demand would still be served.
- *Key corridors and area's of interest are identified based on synthetic truck traffic flows*
  Mapping the demand for hydrogen across the distributed HRS reveals the most interesting areas for the hydrogen mobility market.

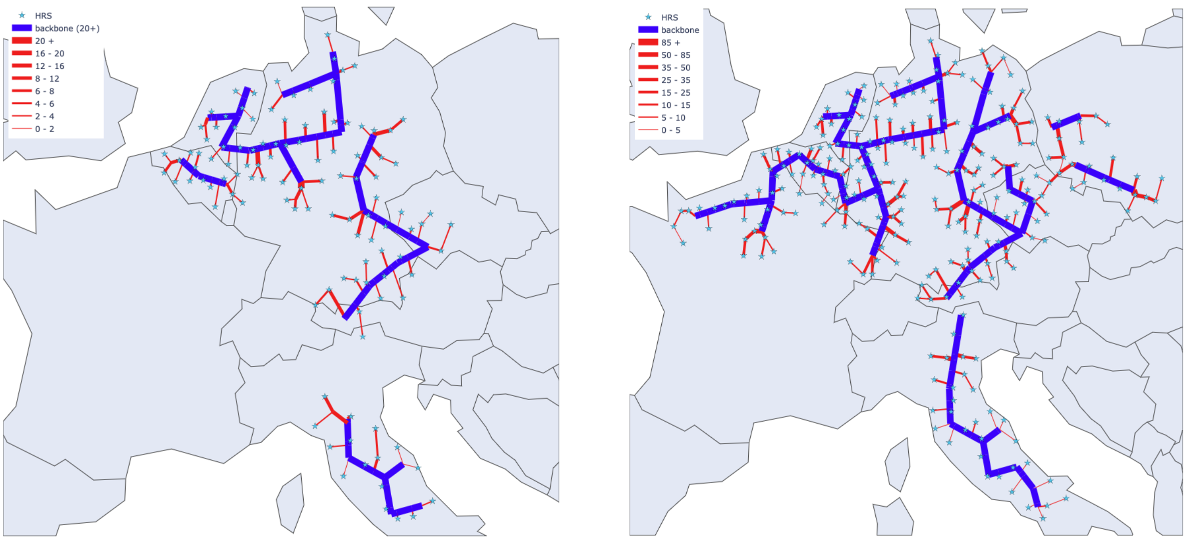
It should be noted that the biases of diesel data this study attempts to avoid are not present in the resulting station topology. The station topology takes into account all performance measures but cannot account for the CAPEX caused by supply. The analysis of the network system design will provide more insight into this.

12.1.4. NETWORK SYSTEM DESIGN

Based on the station location model, the network system design resulted in a hydrogen backbone per decade connected to their respective HRS (chapter 10). Through the work of the optimal network layout tool, this system of pipelines is optimized in terms of costs of connection. The goal is to reduce CAPEX of supply to balance the costs of supply and demand. The networks leave two analyses that form a basis for recommendations; (1) analysis of the distribution of pipelines; and (2) analysis of the share of demand that can be covered by the backbone. The first will shed light on the CAPEX performance measure and the second on the coverage performance measure. The utilization and average distance to an HRS are not relevant for this analysis.

Network system design analysis: pipeline distribution

The optimized distribution of pipelines enables analysis on three fronts; (1) capacities; (2) lengths; and (3) total costs. The latter being determined by the first two. However, prior to this analysis the system is visually analyzed. To this end, the initial results are modified in their format. The width of the backbone is displayed, not linearly but with a predetermined range. This range was determined so that visual differences can be noted on different scales. The legend shows what values are used, note that there is no specific step size.



(a) Demand-based optimal network design 2040

(b) Optimal network 2040

Figure 12.10: Demand-based optimal network design 2040

Visual inspection
As the Optimal Network Layout Tool optimizes the pipeline connections, the topology of the pipelines are of value. Visual inspection of figure 13.5 shows that there is a low degree of interconnection for 2030 and 2040. Interconnection occurs when an HRS is connected to the backbone through another station. A high degree of interconnection leads to relatively low supply costs, since capacity can be shared. On the other hand, it leads to less redundancy. It is therefore interesting for recommendations. Most stations are directly connected to the backbone. As the network design considers stations in close proximity of the backbone (< 100 km) this makes sense. If another station is closer than the backbone, it is more cost-effective to be interconnected. Therefore, larger distances to the backbone would make the network more dependent on other stations. The degree of interconnection may also increase when more stations are built. This increases the likelihood of a station being closer than the supply point on the backbone.

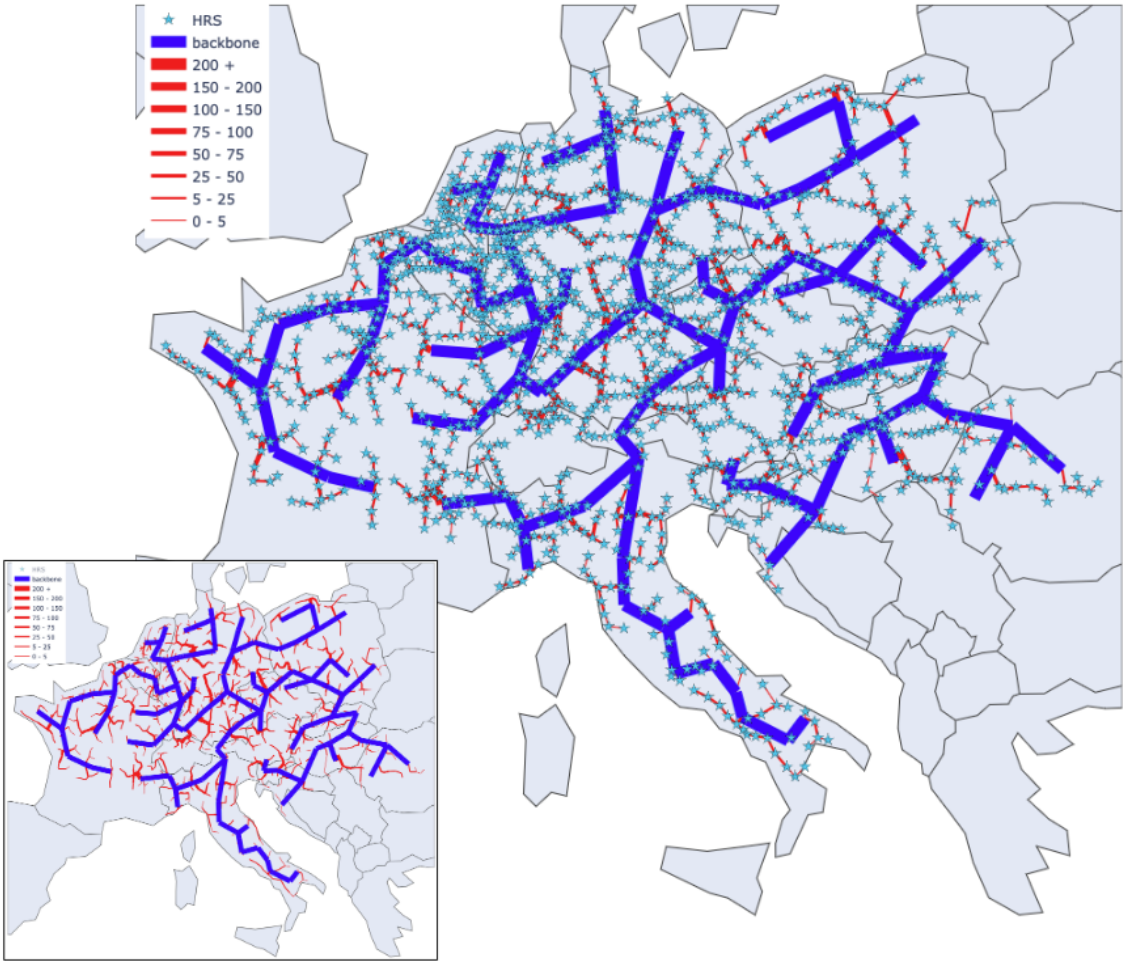


Figure 12.11: Demand-based optimal network design 2050 and smaller visualisation of the network without HRS

This effect is shown by the network of 2050. Interconnectivity sharply increases as many more stations are being built. Demand increases over time, while the average size of a station is not changing. For 2040, the increase is not large, as relatively few additional stations are being built. An event that can be explained by the assumption that the average station size triples during that decade (to 6 tonnes/day by 2040).

Capacities and lengths
The Optimal Network Layout Tool results in a network system model that is able to give capacities and length per pipe as outputs. The results for each decade are presented in the table below:

Table 12.1: Quantification of network system designs

Pipeline extension		2030	2040	2050
Capacity (tonne / day)	*max.*	20.8	89.8	398.0
	min.	1.24	0.55	1.06
	avg.	5.43	16.0	30.0
	total	667482	4384456	41163371
Length (km)	*max.*	95.2	90.5	76.9
	min.	0.16	0.24	0.01
	avg.	34.2	32.4	20.45
	total	4210	8889	28069

There seem to be few real standouts in the results. Maximum, average, and total capacity increase fairly linear over time. One notable result is the decline in maximum and average pipeline lengths in 2050. The maximum length shows that this is caused by the higher degree of interconnectivity. Whereas 2030 and 2040 have a maximum distance close to the max allowed proximity of the backbone (100 km), the 2050 maximum does not come close. For each HRS, there is always another station within 76 kilometers that is connected to the backbone. As stated, this is reflected back in the average length of a pipeline, which is sharply reduced to 20 km by 2050. The total length of piping sees a huge increase due to the increase in number of stations. However, when calculated to tonnes transported per kilometer, the efficiency of the 2050 network becomes clear. The total network averages 1466 tonnes per day per km, where 2040 and 2030 average 493 and 159 respectively.

Total costs of the networks

The Optimal Network Layout Tool also has a built-in cost function, based on which the total costs of the network can be calculated (Appendix G). These calculations are performed and displayed in table 12.2

Table 12.2: Oversimplified costs of supply

	2030	2040	2050
Total costs of supply (Bn €)	3.50	13.8	59.2
Kg H2 transported (t/d)	502	3298	13623
€ / kg (over 50 years)	0.38	0.23	0.24

In addition to the total cost of the backbone, an oversimplified calculation of the levelized cost of hydrogen transportation is included. The total amount of hydrogen transported by pipeline per day is multiplied by the number of days in 50 years. In this way, a sloppy projection of the total amount of hydrogen transported over 50 years (depreciation period) is calculated. It should be noted that this is an extremely unsubtle approach, since (for example) it does not take into account OPEX, annuity or increase in hydrogen demand. However, since this is the case for all three decades, this approach allows a comparison of the different designs.

What is immediately striking is that the cost per kilogram decreases between 2030 and 2040, but stagnates from 2040 to 2050. Further decrease of costs would be expected, as the degree of interconnection would lead to lower costs per kg supplied. However, this effect is offset by the average size of the stations. The larger average station size in 2040 means that relatively fewer stations will need to be built to serve an increasing number of customers. A smaller number of stations, means less pipeline for supply. An effect that is also seen in total length of piping in table 12.1. Note that costs of capacity increase per HRS might than reside in other costs that could inflate the levelized costs.

Network system design analysis: backbone coverage

The last analysis to be performed regarding the design of the network system couples back to the performance measure of demand: coverage. By excluding stations that fall "out of range" of the backbone (100+ km), it can be analyzed how much of demand could be supplied by the network connected to the backbone. With respect to the demand covered by the network of the station location model (> 1 t/d) the following demand coverage results:

- 2030: 28.2% of demand (114 stations, 2 t/d avg.)
- 2040: 43.4% of demand (272 stations, 6 t/d avg.)
- 2050: 65.4% of demand (1372 stations, 6 t/d avg.)

As could be expected, supply coverage grows over the decades as the assumed backbone grows. What is interesting is that the growth is not linear. A large step is taken in the last decade, bringing coverage to over 65% by 2050. This can be explained by the increase in average station size from 2030 to 2040 while it remains the same from 2040 to 2050. The average station size was modeled to increase from 2 to 6 t/d, which causes less stations to be built while serving more demand. This results in a relatively low increase in coverage.

The increase in average station size was modeled to see its effect and it is obvious. Demand-side performance measures do not benefit from this. Relative network coverage is declining, but CAPEX fell sharply. Station size shows to be a factor of influence on both sides of the performance measures.

KEY FINDINGS NETWORK SYSTEM DESIGN

The analyses of the network system designs reveal a number of interesting results that can be taken into account for recommendations. These are presented in an itemzied list for the sake of overview:

- *Average station sizes and degree of interconnection are main drivers of cost reduction*
  The main benefits in terms of cost reduction seem to be in the corner of HRS interconnection. A reduction in the average and maximum size of pipelines can be achieved this way, while the cost of pipelines are also shared (would also include OPEX). Another way to reduce the cost of supply is simply to have less need for supply. Trivial knowledge, but the average size appears to have significant impact.
- *A low levelized cost of hydrogen transportation can be found in pipelines*
  Something that is often raised in the literature is reaffirmed by the results of this study: the cost of transporting hydrogen can become very low when pipelines are used. A very unsubtle method is used, but even these results seem to be consistent with other studies.
- *The growing backbone will increasingly be able to cover a large portion of demand from FC-HDTs*
  Coverage of the backbone over time shows that the pipeline is capable of meeting much of Europe's demand. The 2030 and 2040 backbone would need to be supported by other supply methods. By 2050, however, the backbone is capable of covering all of central Europe via pipeline. This leads to 65% coverage of the demand-based network. It should be noted that the number of stations (and their average size) plays a major role.

12.1.5. TWO-STEP OPTIMIZATION

The two-step optimization aims to find balance between the costs of distance for supply and demand. These costs are determined by the costs for trucks to drive to the station and the costs for piping (supply) to reach the station from the backbone. How this is modeled is displayed in figure 11.1. The aim was to use this optimal network system design of the previous paragraph for the two-step optimization. However, due to optimization problems caused by the interconnectivity of the design this was not possible (section 11.1.3). Therefore, the results of the two-step optimization must be placed into context.

The model is not able to take into account the interconnections of the HRS as supply points. Hence, not every station location can be found that is optimal in terms of supply and demand. However, by searching for the weighted average based on direct connection to the backbone instead of via another HRS, analyses can still be performed. In fact, some of the stations are already doing this. In addition, external factors (e.g., different operators of the stations) can lead to such behavior in practice. The initial application of this two-step optimization revealed the networks as displayed in section 11.2. These networks still enable analysis to (1) discuss the general influence of supply on the station locations, (2) search for costs where (dis)balance would be achieved, (3) review design of the new (sub-optimal) topology.

Two-step optimization analysis: general influence of supply

The topology after the two-step optimization basically reflects the topology of the station-location model, only then the stations are moved towards the backbone. Since direct connections are assumed in a straight line to the backbone, the only degree of freedom here is the degree to which they move on this straight line. To get an idea of how they move, figure 13.8 and 13.9 display results zoomed in on the Netherlands.

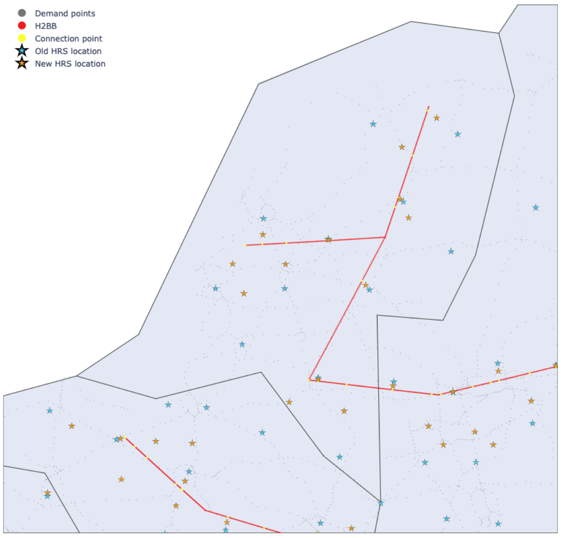


Figure 12.12: Two-step optimization 2030

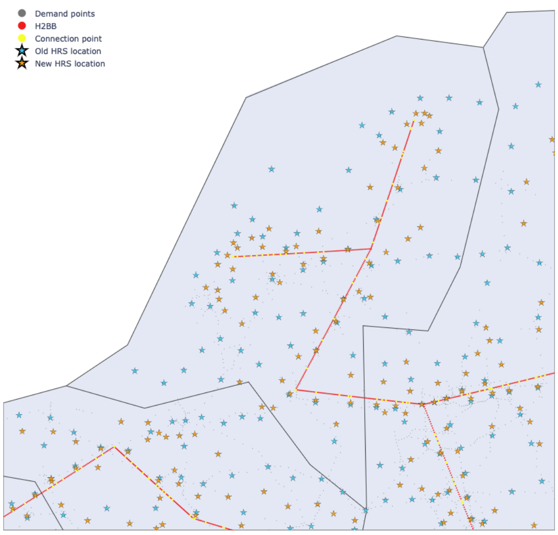


Figure 12.13: Two-step optimization 2050

The results seem to display a balance of supply and demand. HRS do not stay in place, but neither are they placed exactly on the supply line. If one of the costs of supply and demand were to outweigh the other by a factor, it would be clearly noticeable. Moreover, the figures show that some stations are more affected than others. This is caused by distance to the backbone and demand that weigh each HRS.

There are some additional caveats that arise from the handicap of not analyzing the optimal network. If the optimal network were analyzed, the average distance to a supply point would be smaller (caused by the interconnection). As already noted, the effects of supply weigh more heavily as distance increases. Therefore, a less distant relocation of stations would be expected and more realistic. Furthermore, if a direct connection is assumed, no newly constructed pipelines are shared by the HRS. Therefore, there no benefits of sharing capacity are modeled. The optimal network design shows that pipelines are expected to be shared across much of Europe. Therefore, less movement would be expected here as well. Nonetheless, when looking at The Netherlands in 2030, many HRS can be seen that stand alone: they would not be interconnected to reach the backbone. Such locations can be considered optimal, since their position is influenced by their "actual" supply and demand points. It makes sense to consider these results in network planning.

Two-step optimization analysis: balance of costs

To get a better sense of the influence of supply on station locations, calculations can be performed. The total cost figure of supply and demand are considered and it is calculated how much the demand should rise to balance out supply.

- 2030: 2.75 times more demand
- 2040: 1.17 times more demand
- 2050: 3.08 times more demand

It is remarkable that the relative difference in total costs is much smaller for 2040 than 2030 and 2050. This phenomenon is again a consequence of the increase in the average size of stations. As more trucks are assumed to refuel at one station, the impact on demand for that particular station will become relatively much greater.

The figures show that, in 2040, costs overall are almost perfectly balanced. In 2030 and 2050, the costs of supply overall weigh three times more than demand. Keep in mind, that only a portion of total trucking is assumed to be supplied by hydrogen and that no more factors of demand than mobility are taken into account. This might cause the delicate balance that is present, to easily be disturbed. If more FC-HDTs start driving throughout Europe, costs of demand will quickly gain the upper hand. On the other hand, many assumptions have been done to determine the costs of the piping. If one assumption does not hold (e.g. possibility of retrofitting), this might lead to very different outcomes.

Two-step optimization analysis: review of new HRS topology
For the final part of the analysis of the two-step optimization model, the network that results for 2050 is analyzed and compared with the optimal network system design. The choice for the network of 2050 is substantiated by the fact that pipeline analyses involve infrastructure with high CAPEX and low adaptability: it makes sense to consider the distant future in the design. Figure 12.14 presents the station locations that result from both models.

For analysis, the allocation of demand is assumed to be equal, a shift in allocated demand would trigger the interdependecy problems that cause the mathematical complication. In terms of the performance criteria for a network, spatial coverage and total costs of supply remain. Both criteria are analyzed in the same fasion as has been done before. Hence, demand analysis of the station location model (spatial coverage) and supply analysis (network system design) are applied on the two-step optimal network design.

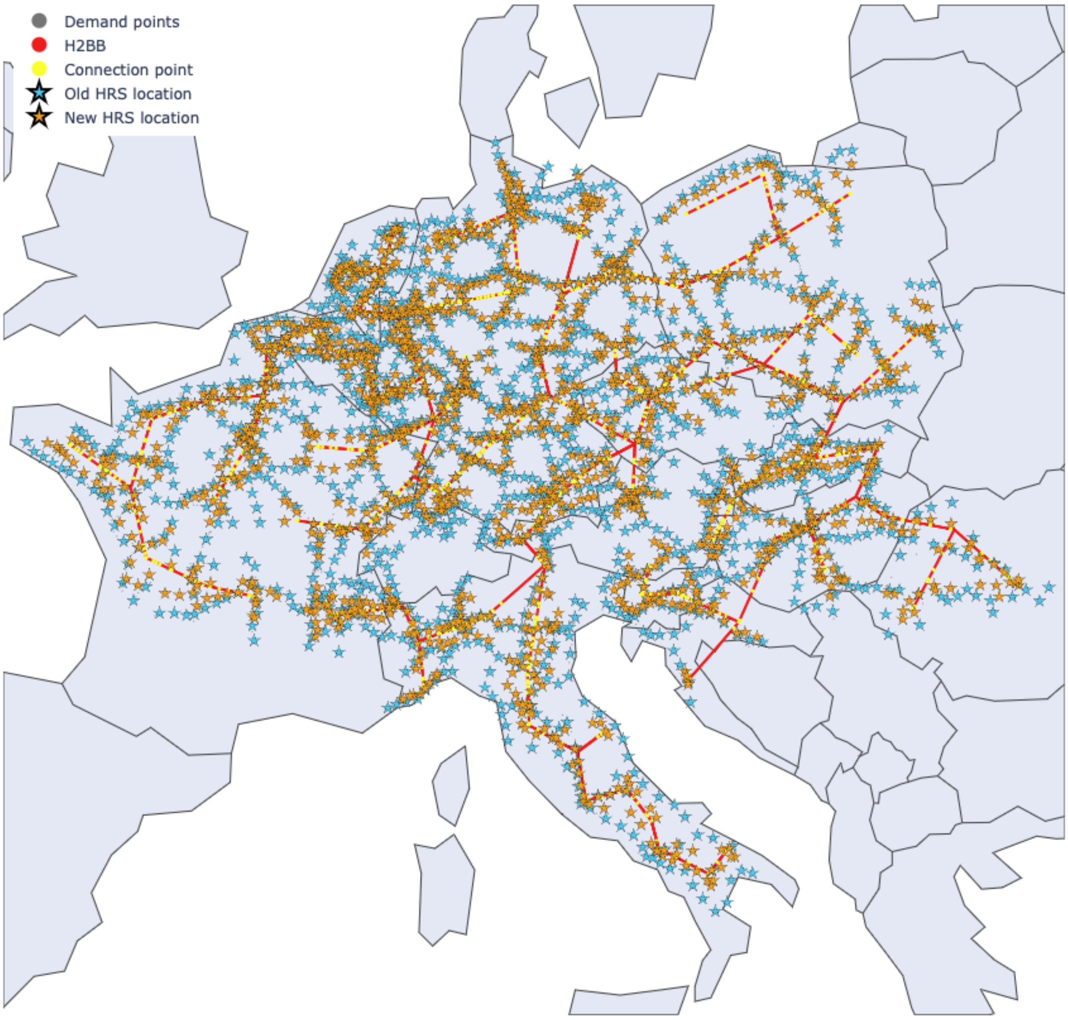


Figure 12.14: Two-step optimization 2050

Analysis of demand: spatial coverage of stations
The two-step optimization relocated the stations to be based on a balance of supply and demand. The demand-based network searched for minimal (weighted) distance from each demand point to an HRS. To quantify performance of the both models, the distance from each demand node to an HRS is calculated. These results are presented in the figures below:

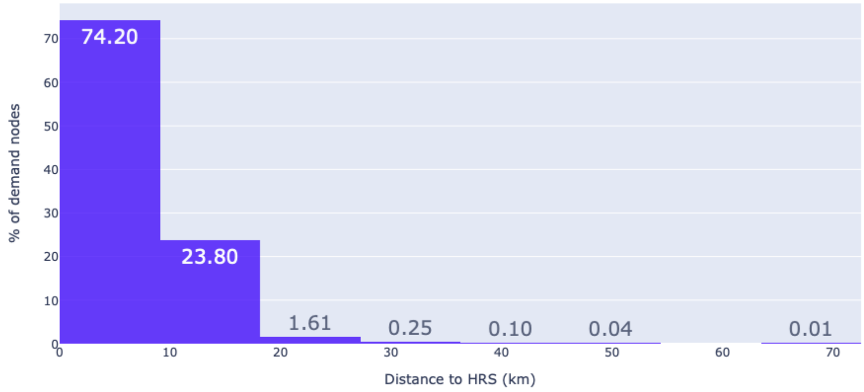


Figure 12.15: Distribution of distance from demand node to HRS: demand network 2050

The figure shows that the demand-based network covers demand very well: 98% of all nodes are served within a Euclidean distance of 20 km, with 74% even within a distance of 10 km. The figure below shows the same results for the network that resulted out of the two-step optimization. It shares the same basis, but locations are moved towards the supply points on the backbone to reduce costs of supply.

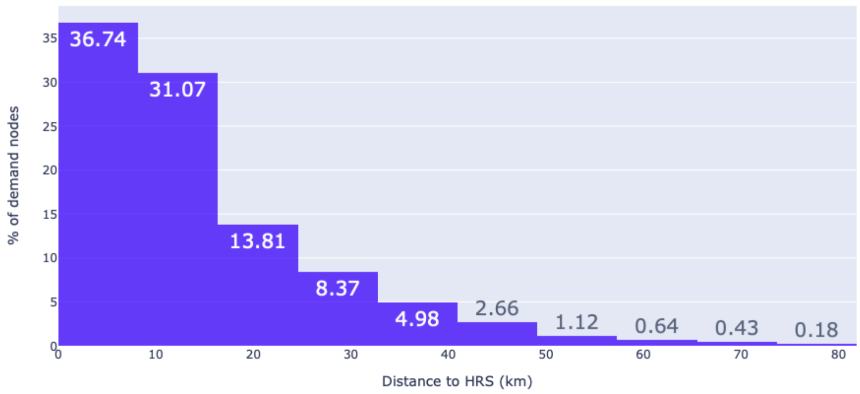


Figure 12.16: Distribution of distance from demand node to HRS: two-step optimized network 2050

The distribution of distances shows a clear decrease in spatial coverage. The distances become slightly more evenly distributed and drivers (on average) have to cover 1.9 times the distance they previously had to, to refuel. This increase leads to a yearly increase of 173.7M€ in costs for demand. Nevertheless, almost 90% of all demand nodes still fall within 30 km range of a refueling station.

Analysis of supply: total costs
The network system design as output of the Optimal Network Layout Tool is presented in figure 14.3. It is immediately noticeable that the total piping required to supply the station is less. This appears to be due to the fact that the stations are placed closer to the backbone. The results are quantified in the table below.

Table 12.3: Quantification of network system designs (incl. optimized)

Pipeline extension		**2030**	**2040**	**2050**	**2050-opt**
Capacity (tonne / day)	*max.*	20.8	89.8	398.0	268.0
	min.	1.24	0.55	1.06	1.00
	avg.	5.43	16.0	30.0	27.4
	total	667482	4384456	41163371	37511559
Length (km)	*max.*	95.2	90.5	76.9	56.5
	min.	0.16	0.24	0.01	0.00
	avg.	34.2	32.4	20.45	14.7
	total	4210	8889	28069	20214

The average pipeline length decreases to 14.7 km, while the maximum pipeline is only 56.5 km long. However, the most interesting results are found on the capacity side. The total, maximum, minimum and average capacities are reduced compared to the 2050 demand-based design. This can be clarified by the minimum pipeline length, which is 0 (not rounded). Due to the two-step optimization, many stations have been placed directly on the supply line, eliminating the need for piping. This resulted in the following cost figures:

Table 12.4: Oversimplified costs of supply (incl. optimized topology)

	2030	**2040**	**2050**	**2050-opt**
Total costs of supply (€ Bn)	3.50	13.8	59.2	43.2
Kg H2 transported (t/d)	502	3298	13623	13623
€ / kg (over 50 years)	0.38	0.23	0.24	0.17

As mentioned earlier, the figure below shows the new design of the network system resulting from the two-step optimization. Recall that the station locations are initially based on the station location model and thus take into account the demand performance measures.

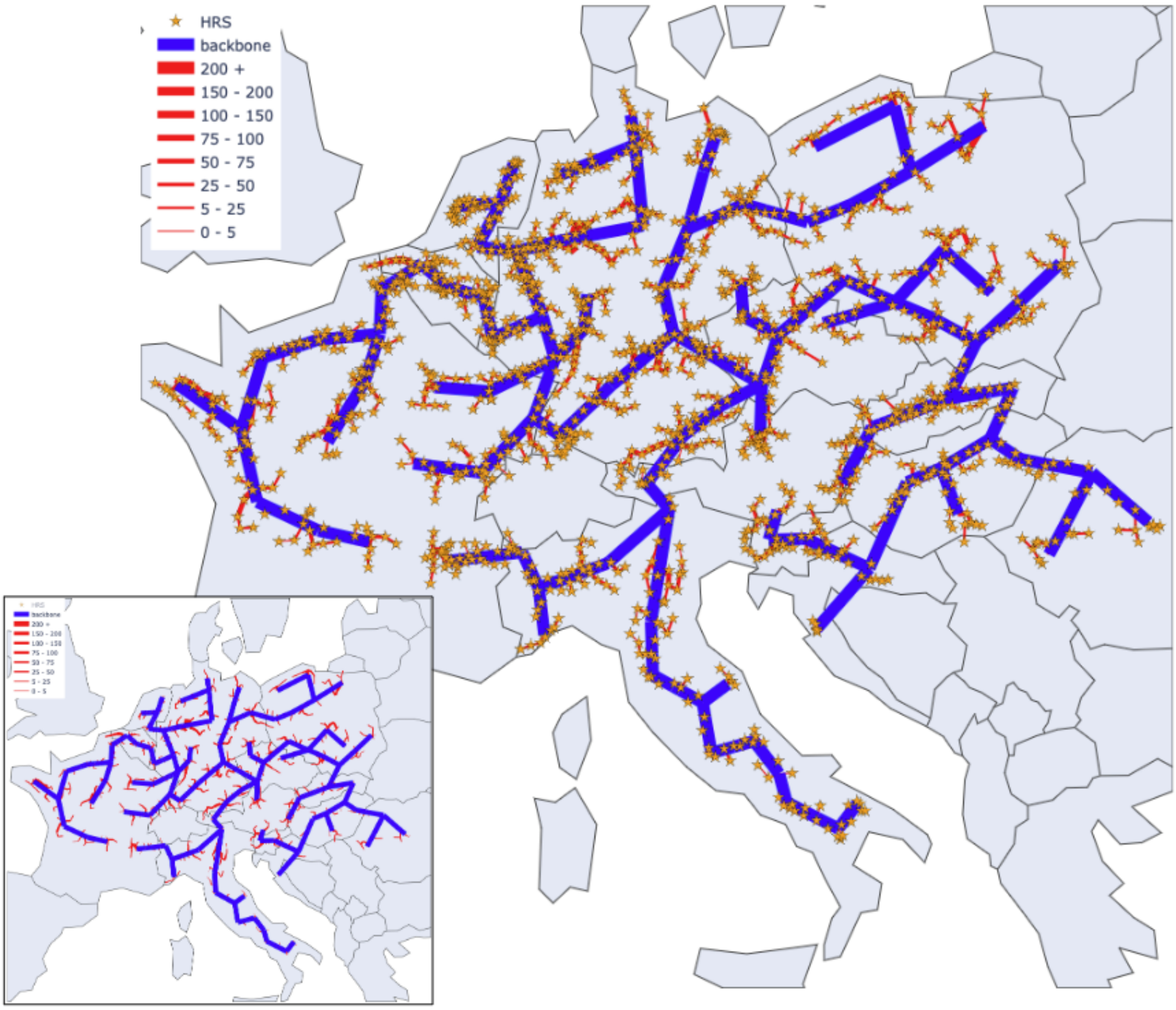


Figure 12.17: Network system design of two-step optimization

KEY FINDINGS TWO-STEP OPTIMIZATION

The analyses of the two-step optimization reveal a number of interesting results that can be taken into account for recommendations. However, keep in mind that not the optimal network is analyzed, but a more simple "direct connection" network. Takouts are presented in an itemized list for the sake of overview:

- *Costs of supply by distribution pipelines are reasonably balanced with costs of demand*
  The model shows a clear balance between the calculated costs of supply and demand. This balance appears largely influenced by station size, as this has high influence on costs of demand: larger stations, means less stations causing trucks to have to drive further. Note that this balance could easily be disturbed by an increase of costs on either side (e.g. higher market penetration).

- *Absolute distance increase for FC-HDTs is relatively low*
  The absolute increase in the distance FC-HDT must travel to refuel remains relatively low. For the optimized HRS location, about 90% still need to drive less than 30 km. However, the total distance increases by a factor of 1.8, leading to a cost increase of 178M€ per year. Note that the demand-based network places 74% of HRS within 10km driving distance their respective trucks.

- *Network optimization in terms of supply has significant effect on hydrogen pricing*
  The network system design of the two-step optimization for 2050 shows a clear reduction in supply costs. In fact, part of the HRS is placed on top of the assumed backbone, practically eliminating supply costs. It should be noted that a decrease in the levelized cost of hydrogen will lead to a cost decrease in the selling price of hydrogen, which may be to the benefit of all parties.

12.2. Underlying assumptions and limitations

In order to reach better conclusions and facilitate discussion, this section will highlight the assumptions and uncertainties underlying each of the modeling steps. Note that the emphasis is on the assumptions and uncertainties in the models. The last chapter of this study will look more closely at some of the models' inputs. The models are discussed in the same order as previously presented.

12.2.1. Freight demand modeling

For freight demand modeling, it was decided to use an economic input-output method. Input and output determinators are used to determine the level of economic activity between regions. In turn, this drives the demand for freight transport. The quantity of goods is ultimately converted to get estimates of the freight transportation demand. This demand for freight transport is finally plotted over a highway nework with the use of a shortest-path algorithm. When using a disaggregated and synthetic tool like this, there are many assumptions that cause uncertainties. There is always the risk that the level of detail is not accurate enough, but some uncertainties are inherent to the assumptions that have been made and these are discussed in this section.

Growth factors 2010 - 2019 - 2030

To calculate the growth factor, a trend-based approach is used and it should be mentioned that this does not reflect sudden changes in trade patterns between NUTS-3 regions. On top of that future regulations could unbalance truck demand.

Empty runs

The factor of 1.25 to account for empty trips is based on the difference between actual freight traffic counts and synthetic data. This is an obvious simplification, since the proportion of empty trips usually varies according to goods and routes. In addition, transport flows within a region are not taken into account.

Routing

As signalled in the methodology, assumptions of routing the vehicles are inherent to the creation of an economic input-output model. Three are identified:

1. The algorithm always chooses the shortest route.
2. Each region is assigned to exactly one network node at which transport routes start and end.
3. If a transport process takes place exclusively within a NUTS-3 region, it cannot be mapped in the highway network.

These assumptions are considered valid, but should be taken into account when interpreting the results.

12.2.2. STATION LOCATION MODELING

As with any model, station location modeling involves assumptions and uncertainties. The analysis of the AFIR locations was based on the allocation of demand to each station, to check for under- and overutilization. Next to that, the coverage of 2 tonnes / day stations were checked. Thereafter, a k-means algorithm was used to create a covering network based on demand. Both forms of modeling have their own shortcomings, with k-means being more standard. However, one assumption is important for both: the distances are measured "as the crow flies". Distances from trucks to HRS are plotted in a straight line, whereas reality would ask them follow a road with all its aspects.

AFIR station locations

The allocation of demand to the AFIR station locations rests upon the freight demand model. So, the assumptions that hold for the database transfer to the allocated demand. The main assumptions of influence are those on the market penetration rates and the fuel-cell efficiency to determine hydrogen demand of a truck. On top of that, the modeling itself brings along it's uncertainties and flaws:

- The station locations itself are still under negotiation making the topic very fluid.
- When analyzing the potential coverage of the AFIR locations, the p-median model could not append datapoints perfectly up to 2 tonnes per day. Therefore, some stations are underutilized and in fact more coverage would be possible.

k-means

Besides the already mentioned assumptions that reside in the database, use of k-means algorithms bring along some uncertainties:

- As stated the algorithm produces different output each time. This leads to slightly different sizes and locations of the stations.
- K-means algorithms have the disadvantage of pre-defining a number of stations. To deal with this, this study decided to determine the number of stations based on the total demand divided by an average station size.

These two drawbacks have already been addressed in the methodology. They do not necessarily conflict with the research objective. However, when interpreting the results it should be kept frond of mind. One should not strive to base network planning on the exact locations that are proposed. As mentioned, both locations and sizes vary slightly each time. As a result, a location by itself is not necessarily optimal. The overall figures provide more of a solid indication of where demand is and enable high-level takeouts on the market.

12.2.3. NETWORK SYSTEM DESIGN

When making use of the Optimal Network Layout Tool, Heijnen (2022) point to a list of assumptions and model limitations [123]. For this study, the following are relevant:

- The algorithm looks for a minimum-cost network without any redundancy. This causes the final and all intermediate results to have a tree (or forest)-topology.
- At each time step, total supply should exactly satisfy total demand.

The lack of a commitment to redundancy is something that would not be desirable in reality. While minimization of costs is certainly important, redundancy plays the opposing role in network planning. However, redundancy is rather expensive and only when the risks are great it will be incorporated. Hence, it is assumed redundancy might only be accounted for within the backbone. This falls out of scope of this study. For refueling stations only, there is little chance it will be accounted for. The consequences of requiring supply and demand to be equal at each time step are already being dealt with in this study. Appendix F shows how the model equals demand and supply. In reality, the backbone can be expected to cover a surplus of supply since it is also used for many other purposes. However, supply is beyond the scope of this study and therefore this limitation is not a barrier.

12.2.4. Two-step optimization

The main uncertainties in the two-step optimization lie in the cost balance. As previously discussed, this balance leaves little room for deviation. If either one of the costs is estimated too high, this will influence results greatly. In general, these costs are hard to estimate and based on many assumptions by itself. Therefore, often unsatisfactory solutions result in these type of optimization. In practice, a highly precise description of the actual goal of the problem is needed. In fact, these should often be described by a globalizing or utility function that combines all the individual distances into a single global value. The other limitation of the model is already extensively discussed: the inability to take into account the optimal network design.

12.3. Validation

Before analyzing scenarios and the sensitivity of the inputs that led to the results, validation is sought. Certain inputs can be benchmarked quite well, which reduces the need for a sensitivity check or a scenario analysis. Moreover, it is necessary to strengthen the foundation of the results presented. Therefore, proper validation is sought and discussed of the data that form the basis of the study: the freight demand and station location model. In this way, an attempt is made to address essential uncertainties.

12.3.1. Freight demand modeling

As this model functions as the basis of the entire study, it is important that it is propely validated. Since synthetic data is created based on an economic input-output model, certain assumptions are inherent. The truck flows are based on real data that is composed of the origin and destination of goods per year. This is translated into a number of trucks that go from A to B over a year. To get to the trucking flows, a large assumption is needed on how the trucks go from A to B. The model used an algorithm that assumes that the trucks use the shortest route. However, this may have shortcomings since trucks may take many different routes based on (e.g.) tolls or traffic. Models based on truck traffic counts would disregard this flaw. Therefore, to give an impression of the data quality, a comparison of their data set has been made by Speth *et al.* (2022) with German truck traffic counts.

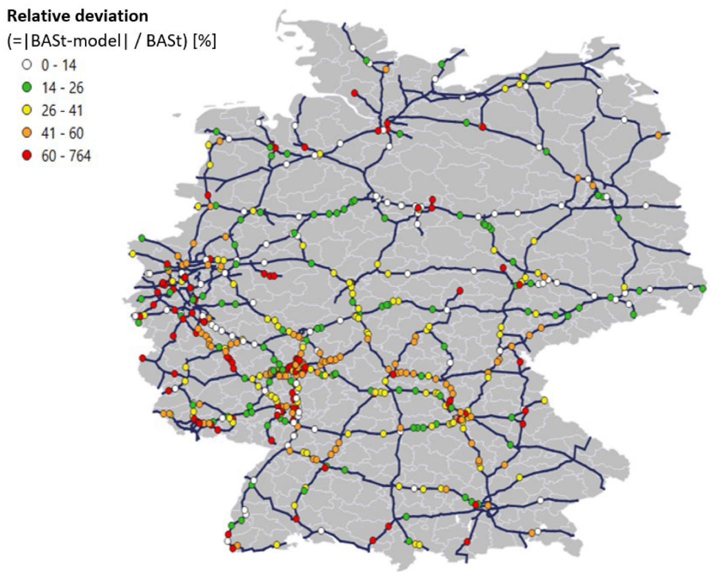


Figure 12.18: Relative deviance of synthetic data to truck traffic counts [10]

Conclusion

The figure shows that on long haul routes the dataset reaches a high degree of consistency. However, in urban areas the consistency gets lower. This can be explained due to the lack of intra regional transport in the model, but might also be due to a less consistent choice for the shortest route in densely populated areas. It can be concluded that for the modeling purposes of this thesis in specific, the data is a very good fit: the HDTs targeted for demand modeling will mainly be used for long haul transport and thus well represented in the data.

12.3.2. Station location modeling

The station location modeling is based on a weighted k-means algorithm. As previously discussed, the major drawback is the disadvantage of pre-defining a number of stations. The number of stations of this study are based on the total expected demand divided by an average station size. This way, perfect coverage can be achieved. The average station size of 2030 was assumed to be 2 tonnes / day, which increased to 6 tonnes / day for 2040 and 2050. This resulted in the following stations per decade:

Table 12.5: Number of stations per decade

	2030	**2040**	**2050**
# stations	964	1285	3527

Since this assumption is at the absolute base of all models, it is important to validate it. The expected demand is validated in the previous section, by comparing the flows to real truck traffic counts. These traffic counts are multiplied by an assumed market penetration that has been previously validated by comparison with figures from a literature review. This leaves the assumption of the 2 tonnes /day open for validation. To this end, estimates of the necessary number of stations are sought and AFIR numbers are discussed.

Estimates of necessary HRS: ACEA

To validate the expected number of hydrogen stations, a study can be sought that specifies the number specifically for fuel cell heavy-duty trucks. A FC-HDT specific ACEA position paper reports on the need for refueling infrastructure specifically for trucking [12]. They note that it is almost completely missing today and that heavy-duty vehicles simply cannot use existing fuel-cell infrastructure for cars. They urge policymakers to take action to ensure a rapid infrastructure roll-out as part of the review of the Alternative Fuels Infrastructure Directive (AFID).

In this appeal to policy makers, they call for 1,000 HRS by 2030 to meet the expected demand. A hydrogen refueling station for trucks should have a daily capacity of at least six tonnes of H2 with a minimum of two dispensers per station.

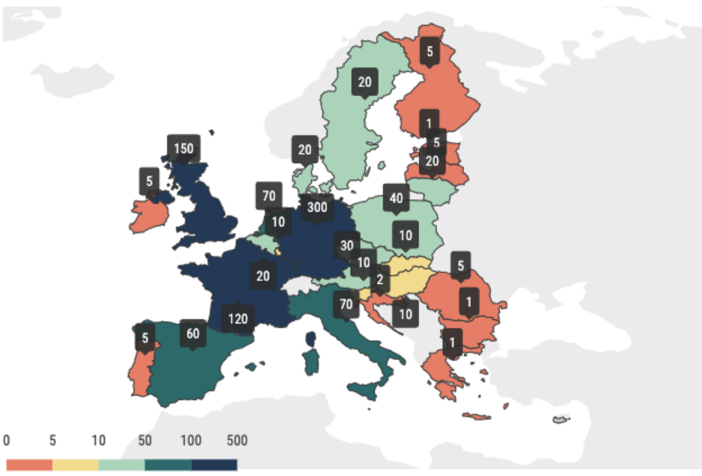


Figure 12.19: ACEA HRS estimates 2030 [12]

The target of 1000 stations is very close to the result of this thesis (964). However, the difference lies in the size of the stations, where they propose a 6 tonne/day station as opposed to the 2 tonne/day of this study. Added to this is the fact that the initial network of 964 stations of this thesis aimed for full coverage of the entire European network. Something that this study clearly does not aim to do, as for example only 5 stations are placed in Portugal and 1 in Greece. A distribution that is more comparable to the network that excluded stations with a demand lower than 1 tonne / day.

AFIR (COMPARISON)

ACEA's position may not fully reflect reality, as they use their numbers to lobby the European Commission. Since they are trying to influence AFIR, this section examines what has come through and tries to draw conclusions.

The AFIR, as currently negotiated, proposes:

- an HRS should be present (700 bar, 2t/day) every 150 km along TEN-T corridors
- an HRS should be present in every urban node

An HRS in every urban node and one every 150 km along the TEN-T corridors lead to at least 457 stations (few more might be necessary to exactly cover the 150 km ask). The figures given by AFIR are the result of a discussion between the main players and the EU, a lot of thinking and modeling can be expected to be behind the regulation. This is noticeable in the locations of the "urban nodes". These nodes are clearly not evenly distributed over Europe, and see a much higher density in the areas that have high truck flows. It is no coincidence figure 14.2 shows a relatively good distribution of demand over the AFIR stations. Over- and underutilization, altough present, are within the realms of reality. There seems to be no clear mismatch between the data. As the AFIR proposition stems from a discussion between the government and private parties: the results of AFIR provide regulations that ensure a large share of coverage, but also serve the market well. The design of this thesis that also had this target, excluded underutilized stations with low demand. This lead to 531 HRS in 2030.

Conclusion

A clear comparison with the results of ACEA and AFIR is not possible, since both are used for different purposes. However, the total figures of 1,000 stations from the ACEA and 457 from the AFIR are certainly consistent with the results of this study. The results of this study are presented as the amount of stations prior to the 1 tonne demand threshold and with this threshold in place. Recall that the former was aimed at covering the entire range of demand and the latter is more balanced with the cost (performance measures) of supply. An overview is given in table 12.6.

Table 12.6: Comparison of external sources and this study: number of stations

	External sources		This study	
	ACEA	**AFIR**	**No threshold**	**1 t/d threshold**
Number of stations	1000	457	962	527
Demand	6 t/d	2 t/d	2 t/d	2 t/d

The ACEA suggests larger stations and about as many as the network without a threshold from this study. The network without a threshold (12.6) is even somewhat in line with what figure 12.19 shows. This means that the ACEA suggests some stations to come online in demand infavorable regions. The study with demand threshold shows that these are of relatively low interest. Their reasons for presenting these figures could be to exert subsidiary pressure on governments. The AFIR is more the result of a discussion between players with different objectives. Their result of 457 two tonne per day stations located at urban nodes should be based on a balance. This thesis represents this balance by the station topology with a 1 tonne demand threshold. These numbers are very consistent with the station locations of AFIR. Both these comparisons with ACEA and AFIR validate the input of the number of stations for this thesis. It can even be argues why there are differences between ACEA and AFIR based on the study results. However, there is one main difference which reveals the value of this thesis: the topology. The topology of the urban nodes in AFIR is quite different from that of the network proposed in this thesis. The network of this thesis focuses on a balance between supply and demand, where that of AFIR stems from a political point of view. This point of view takes into account balance, but could lean too much towards covering Europe, potentially creating an unfeasible network. The value of this study lies in the arguments that show a more balanced approach to network planning.

13

Scenarios and Uncertainty

The final chapter, which precedes the conclusions, addresses the final sub-question (presented in the table below). First, a basis is established of what critical uncertainties for the design of a hydrogen refueling network are based on an interview with an expert. Thereafter, these uncertainties are analyzed to see what their influence might be on the results of this thesis. Results are presented in section 12.

Table 13.1: Research Goals of Station Location Modeling

Main Question	
Assuming a European hydrogen backbone, what is the optimal configuration for a European hydrogen refuelling station network for heavy-duty trucks based on expected demand?	
Subquestion addressed in this section	**Research Method**
What are critical uncertainties for the design of a European hydrogen refueling network?	Scenario exploration

13.1. Uncertainties

To uncover uncertainties that are critical to the network design, an interview is conducted with Justin Starreveld (J. Robbins, personal communication, April 8, 2022). Starreveld is currently working on a Phd to support strategic investment decisions regarding the production and deployment of hydrogen at scale. The aim is to support mathematical modeling techniques that determine an optimal transition path. He notes that for such large and complex problems, there is much uncertainty regarding future developments in technology, economics, and policymaking. Therefore, the main goal of his PhD project is to develop a robust optimization methodology for dealing with such uncertainties while optimizing long-term, large-scale energy system models.

In general, Starreveld speaks of three forms of uncertainties for large infrastructures; (1) forecast errors; (2) estimation errors; and (3) implementation errors. His estimates are that these errors will play the largest roles in (1) technologies, (2) politics, and (3) hydrogen pricing. For hydrogen infrastructure in specific, these errors mainly lead to uncertainty in costs of hydrogen techniques and demand for hydrogen. Questions on the future demand of every sector and the development of hydrogen techniques (both within the FC-HDT and the piping) find themselves at the basis of uncertainty and answers would support strategic decision-making. In general, Starreveld states that there are two ways of dealing with uncertainty in modeling:

- stochastic (not too many parameters)
- robust (best reaction to extreme scenario's)

13.1.1. HYDROGEN DEMAND

By excluding all forms of demand except for heavy-duty trucks (section 4.1), this thesis has already taken a first stochastic step. There are few demand parameters left. However, this demand is still considered uncertain because the literature varies widely in terms of the expected market penetration of FC-HDTs A.2. Therefore, scenarios driven by fluctuation in demand are explored for this thesis. As insinuated by Starreveld, this is best done by a best and worst case secnario.

HYDROGEN DEMAND: MARKET PENETRATION RATE

The demand for hydrogen from FC-HDT for this thesis is determined by the underlying dataset and the market penetration rate. The plotted demand is then used for analysis of the AFIR and the station location model. The latter forms the basis of the rest of this study. To test the uncertainty of the input parameter, a best and worst case scenario were analyzed. The best and worst case scenario's can be derived from section 4.1.2. There, scenario's are described that meet climate goals, but see different market outcomes for alternative powertrains.

The AFIR sites are analyzed for a worst-case scenario, as this is most interesting for policy advise. A best-case scenario would only show that stations are less underutilized. That output would have little value for the discussion of regulation in this case. The station location model is only tested in a best case scenario as current input might be on the low side of expectations. The market penetration rate that was selected in section 4.1.2 was the Power to gas dominated scenario. Hydrogen and electric powertrains form a very plausible scenario in the eDrives scenario, as the technology is at a relatively advanced stage.

Worst case scenario: AFIR analysis

For the worst case scenario, the market penetration rate that was previously analyzed (5% in 2030) is lowered by a factor of 5 to 1%. This was done to mimic a very low uptake, but not non-existent. The Power to liquid scenario of section 4.1.2 which yields 0% uptake would not give any new insight. The figure below reveals the level of underutilization over Europe as an effect of a hydrogen market that does not (or barely) take off:

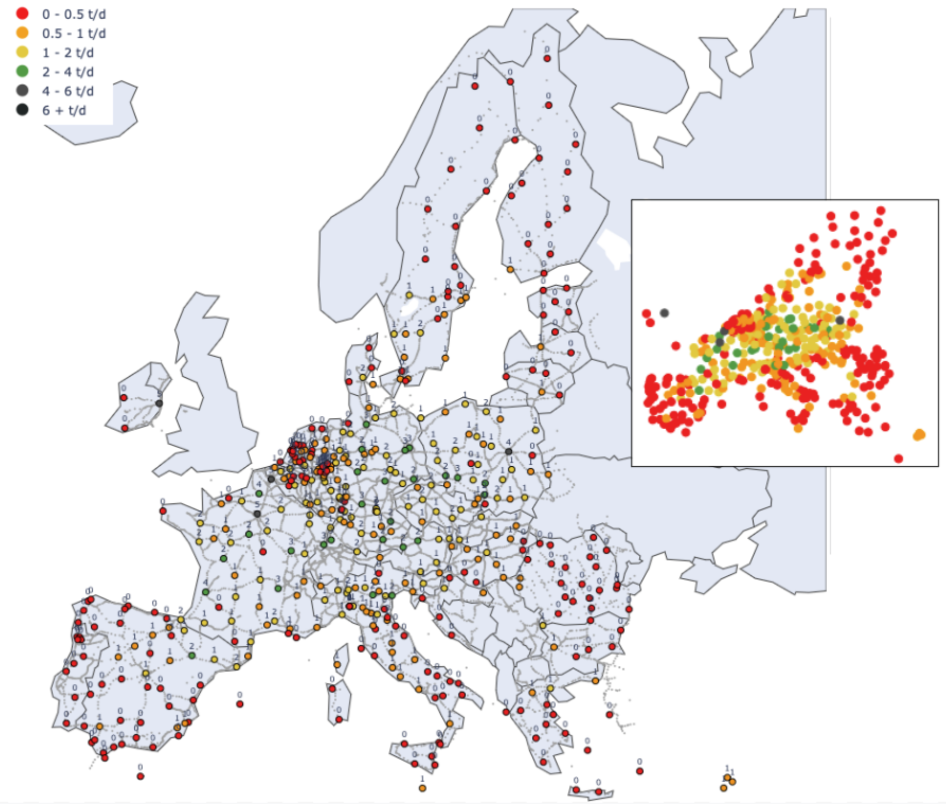


Figure 13.1: AFIR analysis: worst case scenario

The figure shows that all peripheral areas of Europe are severely underutilized, while overutilization is never an issue. Only Central Europe would still be able to achieve a decent demand rate. The analysis shows the importance of penetration in the FC-HDT market. If the predicted market penetration is lower than expected, a huge underutilization with all its financial consequences could be the result of the policy.

Best case scenario: station location model
For the best case scenario, the eDrives scenario from section 4.1.2 is taken as the basis. This scenario predicted FC-HDT market penetration of 15%, 45% and 80% for 2030, 2040 and 2050, respectively. This market penetration yields the following topology for 2030:

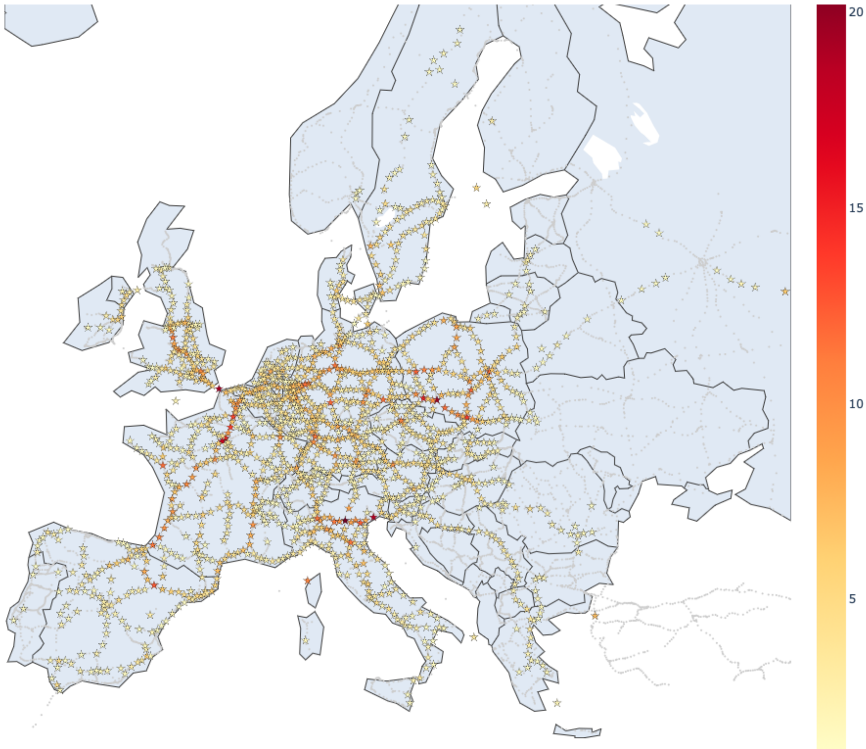


Figure 13.2: Station Location Model: best case scenario market penetration rate (legend shows t/d)

As the average stations size did not increase (2 t/d), the number of stations sprouted. Where the main scenario lead to 529 HRS over 1 t/d being built, this scenario projects 1538. This leads to the following distances from each demand point to an HRS:

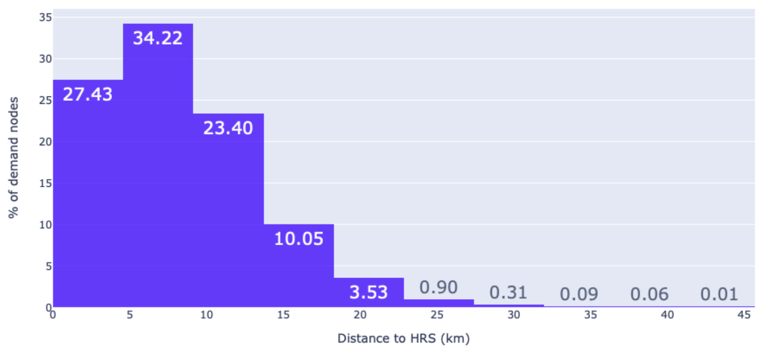


Figure 13.3: Distance to HRS: best case scenario

The distribution shows that 95% of the demand points are reached within 20 km of an HRS. This significant decrease is expected due to the increasing number of stations. What is more interesting is that the distribution has remained exactly the same. The same effect is noticed for the archetype distribution given below:

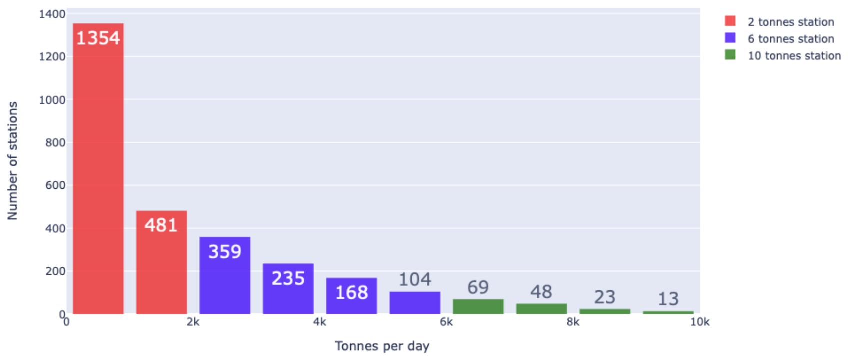


Figure 13.4: Archetype distribution: best case scenario

No matter how different the numbers are, the distributions remain the same. This exposes a property of the model: the distributions of distances and archetypes grow accordingly to the market penetration rate. This is explained by the the number of stations that are pre-defined and based on the average station size: the eDrives scenario and the Power to Liquid scenario.

CONCLUSIONS

A worst and a best case scenario analysis is performed on the analysis of AFIR and the station location model respectively. The AFIR sites are analyzed for a worst-case scenario, as this is most interesting for policy advise. The station location model is only tested in a best case scenario as current numbers might be on the low side of expectations. The scenario analysis for demand revealed importance of the market penetration rate on the AFIR implementation. An incorrect prediction of market penetration could lead to large financial losses due to underutilization. In addition, the best case scenario analysis showed how this particular model responds to uncertainty in demand. Since the number of stations to be modeled is determined by the market penetration rate, the results follow accordingly. The topology of the best case scenario showed that in this scenario many stations of 2 tons/day need to be built by 2030 (almost unattainable).

13.1.2. HYDROGEN SUPPLY

Next to the uncertainty of demand, Starreveld notes that the development of technologies in the hydrogen sector are highly uncertain. For this study, this concerns the FC-HDTs (e.g. efficiency) but also costs that reside in pipeline supply. The cost balance of the two-step optimization and the total costs calculated for the network system rely on costs per kilometer of pipeline. For this study, these pipelines were assumed to be of the distribution type. This yields relatively low costs and the pipes are used for short connection. The pipeline costs itself are pretty well defined. However, it is not explicit when transmission and when distribution pipelines should be placed. The assumption is made that only distribution pipelines would be used. However, transport up to 100 km was possible, which may not be in line with reality. This is reaffirmed by the notion in section 4.2.1 where uncertainty about the use of this single value for piping was present. Therefore, the effects of a change in costs of piping are to be tested by assuming transmission pipelines instead of distribution. Both the network system design and two-step optimization are performed with transmission pipelines.

HYDROGEN SUPPLY: USE OF TRANSMISSION PIPING

Higher costs of piping are not the only aspect that come into play when analyzing transmission pipelines. A new "range" of the backbone must also be considered. Transmission pipelines can span a much greater distance than distribution pipelines, and thus connect more distant HRS to the backbone. However, the HRS may be interconnected, making a threshold value of distance from the backbone itself of less value. Therefore, the maximum distance from previous research (500 km) is taken to be "within the range" of the backbone as determined in section B.3.2.

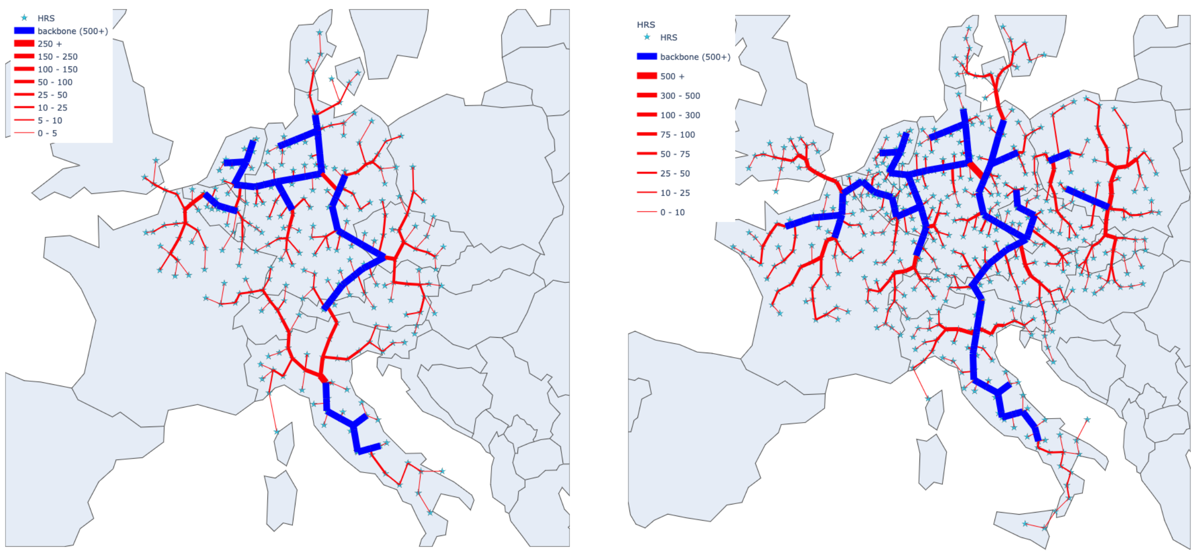
Network system design analysis: backbone coverage

As for the distribution pipeline network, the initial analysis is fed back into the demand analysis. By excluding stations that fall "out of range" of the backbone, the ability of the backbone to supply FC-HDTs in Europe over time is given. With respect to the demand covered by the network of the demand analysis (> 1 t/d) the following supply coverage result:

- 2030: 58.2%
- 2040: 72.7%
- 2050: 81.6%

As could be expected, supply coverage grows over the decades as the assumed backbone grows. Interestingly, the 2030 network is already able to cover well over half of total demand. This coverage is almost double that of the distribution pipeline network. However, the relative increase in coverage decreases over the three decades: by 2050, there is only a 15% increase. What is already evident from this output is that modeling transmission pipelines is more like extending the backbone. The 2040 network already looks more like a full-fledged European backbone than one in a preliminary stage. Hence, analyzing the 2040 network system works better for reasons of comparison. Moreover, the analysis of an early expansion of the backbone also has added value. Therefore, instead of the 2050 network, the 2040 network is chosen for further analysis.

Network system design analysis: pipeline distribution
As with the distribution network, the system is visually analyzed. Results are again modified in their format to fit their output data. The legend shows what values make up the format, note that there is no specific step size. Visual inspection for 2030 and 2040 (13.5) leads to a first remark: the methodology for allocating supply for individual parts of the backbone was not successful. The separate part of each backbone was supposed to be supplied by one source, which would avoid interconnection of the backbones. However, the results of figure 13.5 show that the parts are getting connected. Since this analysis is for transmission pipelines, which are also the basis of the backbone, this is not considered a problem for this analysis. On the contrary, it gives insight into possible extension of the backbone.



(a) Optimal network 2040

(b) Optimal network 2040

Figure 13.5: Optimal network system designs

The connecting of the backbones reveal that the backbones cannot supply their respective demand nodes themselves. A phenomenon caused by the interconnectivity of the HRS linking up to connect to a backbone far from their own location. This complicates the pre-determination of supply. In this line of thought, a preliminary conclusion can be drawn: a complete backbone has relevant impact on security of supply. The connection of the backbones shows the value of an extensive backbone connected to multiple supply sources when demand is uncertain.

The same capacity, length, and cost values that were calculated for the distribution network are displayed in table 13.2 and 13.3 with the addition of the tranmission piping output.

Table 13.2: Quantification of network system designs (+ optimized and transmission piping)

Pipeline extension		2030	2040	2040-tra	2050	2050-opt
Capacity (ton / day)	*max.*	20.8	89.8	323	398.0	268.0
	min.	1.24	0.55	0.02	1.06	1.00
	avg.	5.43	16.0	28.5	30.0	27.4
	total	667482	4384456	15090044	41163371	37511559
Length (km)	*max.*	95.2	90.5	181.5	76.9	56.5
	min.	0.16	0.24	0.50	0.01	0.00
	avg.	34.2	32.4	43.65	20.45	14.7
	total	4210	8889	23137	28069	20214

The transmission pipeline network shows a very large increase in capacity, nearly equal to the projected distribution network in 2050. For length of the pipes, since there is already little assumed backbone in the ground, the added network must cover much more ground. For this reason, the average and maximum lengths are much greater than the 2050 network.

Table 13.3: Oversimplified costs of supply (with transmission and optimization)

	2030	2040	2040-tra	2050	2050-opt
Total costs of supply (€ Bn)	3.50	13.8	312	59.2	43.2
Kg H2 transported (t/d)	502	3298	5661	13623	13623
€ / kg (over 50 years)	0.38	0.23	3.10	0.24	0.17

The 2040 transmission network shows a huge increase in the total cost of supply. This is easily explained, since a complete expansion of the backbone will be built solely for the purpose of mobility. The major difference is in the properties of the pipes (e.g. steel/plastic). In addition, the construction of new pipelines is assumed where retrofitting would lead to a large cost reduction. This was assumed because it would be an addition to the backbone that would already be in place. The two-step optimization will take a closer look at this imbalance.

TWO-STEP OPTIMIZATON

Two-step optimization analysis: influence of supply

As for the distribution network, the optimal design of the network system could not be analyzed. A network with direct connection to the backbone will again be used to perform the same analyses. Figures 13.6 and 13.7 display the locations based on the weighted average of the demand points and the backbone connection:

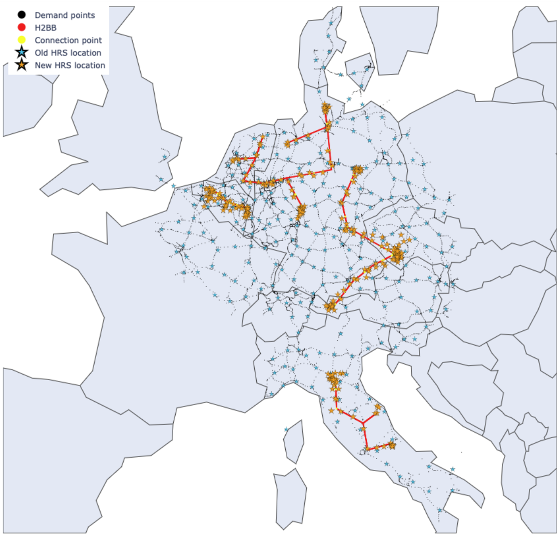


Figure 13.6: Two-step optimization 2030

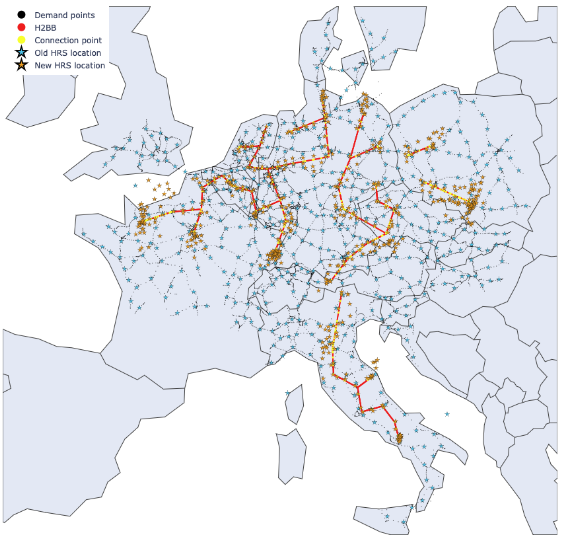


Figure 13.7: Two-step optimization 2040

As an initial inspection of the figures immediately makes clear, there seems to be no balance in the cost of supply and demand. The zoomed in results of the optimization for 2030 and 2040 (13.8 and 13.9) show the HRS get placed almost exactly on the line of supply. The costs of supply by backbone heavily outweigh the costs of distance to demand. If the design of the network system were optimized, the impact of this imbalance would be less noticeable. Capacities would be shared and average distances to supply points would be shorter. Nonetheless, the idea of balance is far from the case.

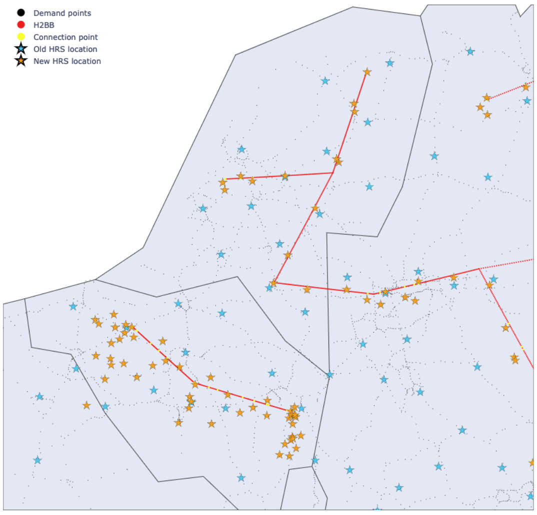


Figure 13.8: Two-step optimization 2030

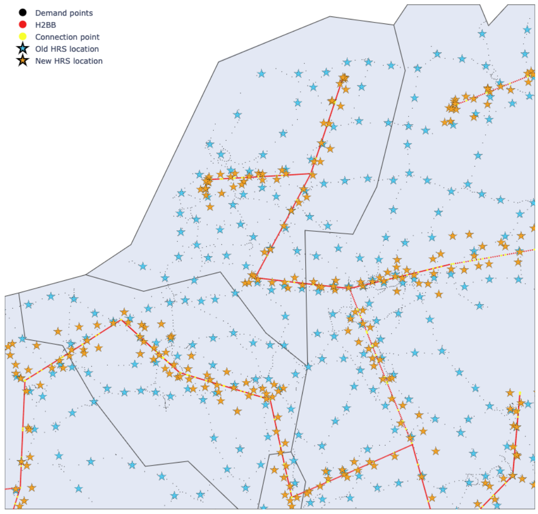


Figure 13.9: Two-step optimization 2040

Two-step optimization analysis: balance of costs

To get a better grip on the disbalance of transmission piping and station location determination, values are sought to find balance. Both total cost figure are considered and it is calculated how much the demand should rise to balance out supply.

- 2030: 55 times more demand
- 2040: 24 times more demand
- 2050: 36 times more demand

The same effects previously seen for 2040 are present. The average station size influence the balance of costs. The numbers show a huge increase in demand to get to a balance between supply and demand. This makes clear transmission piping for delivery all the way up to an HRS does not make sense. The disproportional ask to balance out costs are already clarified: the model assumes that only FC-HDTs would use this extremely large supply network of strenghtened steel, high capacity pipes.

CONCLUSIONS

The use of transmission lines to build a supply network for the HRS proved to be an outright mismatch. Another look at the optimal network design (13.5a) shows the scope of the newly built piping for the HRS. A huge extension of the backbone is modeled, mimicking the extension of a backbone only to supply HRS. Note that the backbone itself was already modeled based on expected demand from mobility. On top of that, the backbone was modeled with other demand in the mix as well (e.g. industries). To ask for such a large expansion, solely for mobility, and in the year 2040, is disproportionate.

These conclusions reaffirm the choice for distribution piping, but also show the sensitivity towards a price increase in piping. No extension of the assumed backbone is desirable, it seems to serve as a solid foundation to build on. This is reaffirmed by the analysis that coverage of the distribution network catches up with the transmission network over time. The 2040 transmission network resembles the 2050 backbone with distribution pipelines added to it. An interesting (but obvious) finding that also emerged was that, when supply is uncertain, a connected backbone adds value.

14

Conclusions, Discussion Recommendations

This report is written with the aim to support Shell in their decision making for network planning of hydrogen refueling stations over the course of the next three decades. Shell's plans to date have been based primarily on diesel data from their own stations. Their analysis therefore include three forms of bias; (1) cross-border fueling due to varying excise taxes; (2) data gaps due to Shell's low presence in certain countries; (3) little precision about what part of the fueling is done by heavy trucks. In response, this thesis has the objective to investigate the optimal topology of a network of hydrogen refueling stations for heavy-duty trucks that takes these biases into account. The intended outcomes were recommendations and insights for station locations, along with their size. Note that, to achieve this, not only the demand but also the supply of hydrogen is considered. The cost of supply cannot be neglected, as a balance must be present between the cost of supply and that of demand.

Summary of research approach

To fulfill the research objective and cover identified research gaps, a two-step optimization approach is implemented. Prior to this, the literature surrounding freight demand modeling was consulted. This lead to the choice for an economic input-output model that enabled data analysis on a European level. Using this data, a weighted k-means algorithm was used to clusterize a pre-defined number of stations based on demand. This second step was the first optimization, which provided the first insights into the market. The third step was the modeling of the network system. By assuming a hydrogen backbone, and placing that in the same plain as the demand-based network, use of an optimization tool was made possible. This Optimal Network Layout Tool connected each station to the assumed backbone in optimal fashion. This released output for further analysis of cost and network design, yielding important findings. Finally, the goal was to re-optimize the demand-based network with supply influence. A balance between the two costs would be achieved, completing the two-step optimization. However, the interdependency of supply points in the network design stood in the way as HRS get interconnected. Nevertheless, insights were gained by assuming a direct connection to the backbone.

Following the analysis of the results, the limitations and assumptions embedded in the models were discussed. In addition, an attempt was made to validate the fundamental modeling steps of the study. Finally, uncertainty in (hydrogen) infrastructure models is discussed and ways to deal with it were presented. Two uncertainties of this study were addressed in a scenario analysis to reveal how the model responds. Next to uncovering model behaviour, more outcomes for recommendations did emerge.

This chapter will draw conclusions based on the key findings of the result and scenario analyses. Thereafter, the conclusions will be discussed in terms of their vulnerabilities and strengths. Finally, clear recommendations for Shell will be made.

14.1. CONCLUSIONS

To fulfill the research objective, this study aimed to answer the following research question and contribute to the network planning of HRS:

Assuming supply from a European hydrogen backbone, what is the optimal configuration for a European hydrogen refueling station network for heavy-duty trucks based on expected demand?

The answer to this question might seem to steer toward a single design that would be the end result. However, this thesis adopted an approach of searching for this design while collecting key outcomes along the way. These key findings lead to recommendations that are the real value, since in reality no design is perfect for network planning. To structure this path, sub-questions were created. The conclusions are presented in the same manner.

Prior to analysis, performance measures of other studies were collected and lead to the following set of criteria for a good network; (1) utilization rate of HRS; (2) the weighted average distance (travel time) to an HRS; (3) to maximize total vehicle kilometers covered; and (4) to minimize CAPEX (or number of stations). These criteria are not tested directly for each network, but serve as a guide to the results.

What are relevant techno-economic parameters for a hydrogen refueling station network and its supply?

The search for relevant techno-economic parameters culminated in a full system description in chapter 4. The section functions as an exploration of what is out there and what is relevant for this study. EU category N3 and O4 were taken as the scope for FC-HDTs. Their characteristics in size and weights are atuned to this categorization. Next to that, the range of these vehicles as well as their expected market penetrations are relevant for network planning. The range is set at 800 or 1445 km, depending on the fuel storage (350 or 700 bar, with 50 or 85 kg H2). Market penetration rates are expected to be 5, 15 and 30% for 2030, 2040 and 2050, respectively. Furthermore, two hydrogen supply chains were analyzed. Proper delineation followed, where decisions were made about which input values and cost figures were considered appropriate. These were informed decisions based on the literature and current practices of hydrogen in heavy-duty mobility and the associated refueling infrastructure. For example, the price of hydrogen at the pump was set at 4 euros per kg, for all three decades, and a distinction was made between distribution and transmission lines. The latter is of critical influence to network planning, as the costs for piping largely determine the costs of supply.

What are the current plans and institutions for hydrogen refueling stations and how do they influence the network design?

The second sub-question seeks to identify the political boundaries within which the network should operate. This provided opportunities and threats for the hydrogen market, and for mobility in particular. On top of that, two regulatory constraints from the AFIR emerged that are highly relevant to network planning. The Alternative Fuels and Infrastructure Regulation is currently included in the form of a directive (AFID), but is under review. It is set to become a regulation and therefore binding on each member state. The AFIR currently prescribes that: an HRS should be present (700 bar, 2t/day) every 150 km along TEN-T corridors and that an HRS should be present in every urban node. Once negotiations are finalized, the basis of any HRS network should be the constraints inflicted by the AFIR. However, as the regulation is still under negotiation, analysis of the proposed station sites was deemed more useful. Results can be used to inform the negotiations.

To come up with clear arguments for recommendations, the AFIR is analyzed on two topics: (1) potential under- or overutilization of the mandatory station locations and (2) their coverage of demand in Europe. This analysis resulted in the following key findings:

- *The AFIR determined HRS in peripheral areas of Europe are headed for underutilization.*
- *AFIR stations in central Europe do not nearly cover all demand at 2 tonnes per day.*
- *The addition of an HRS every 150 km, on top of the proposed urban nodes seems unnecessary.*

Support for these conclusions can be found in section 12.1.3. The results show a clear image of underutilization of AFIR HRS at the peripheral areas and overutilization in central Europe. That some station locations, as determined by AFIR, would lead to underutilization must not come as a surprise. As a government body, the European Commission is tasked with ensuring coverage across Europe. Even where the market sees little potential. Nonetheless, if stations are headed for large financial failure, this would not be a reasonable ask. Therefore, the scenario analysis (13) includes a worst case scenario. This worst case scenario assumes a 5-fold drop in demand, which would leave only a few centrally located stations viable. Both these analyses should be used in discussion with the European Union on the regulation. To this end, the last finding is also helpful. Negotiations on AFIR can be about station size or location. The latter is composed of "urban nodes" and a distance threshold. While the distance threshold is easier to negotiate, it has less impact on network planning. The urban nodes seem to be well thought out locations that cover almost all of Europe, hence they should not go undebated.

What is the expected European hydrogen demand from heavy-duty trucks over time, distributed over a highway network?

This sub-question was asked to find a basis for modeling, that would omit the Shell diesel data biases. To this end, a method was sought that would provide data as "truck traffic counts". The freight demand model, based on economic input-ouptut served this purpose well. Note that it was intended primarily as a basis for further analysis, and it functions well that way. Initial results of the freight demand model gave a clear overview of where trucks were located over a year. This initial dataset of traffic flows is converted into data points suitable for the following research methods. Vehicle kilometers per year on each datapoint were calculated and multiplied with the expected market penetration rate. Kilometers that in turn could be recalculated to local demand for kilo's of hydrogen to answer the sub-question. Based on the system description it was assumed a fuel-cell heavy-duty truck can drive 1440 km on 85 kg of hydrogen (at 700 bar). Therefore a FC-HDT can drive 17 km (1440/85) on 1 kg of hydrogen. Hence, 17 vehicle kilometers on a datapoint represent demand for 1 kg of hydrogen.

How are hydrogen refueling stations distributed over Europe in order to optimize coverage of European FC-HDT demand?

The fourth sub-question was posed to find station locations optimal in terms of demand: an optimal demand-based network. The application of the weighted k-means algorithm immediately ensured two out of four criteria to be covered: minimization of the weighted average distance and maximization of vehicle kilometers covered. The demand for each data points got allocated to the respective HRS. This initial topology was further adjusted in the direction of the other criteria. By iteratively raising a threshold to omit stations (from 1 to 6 tons per day), and measuring coverage each time, an optimal balance is found between coverage and the number of stations. By assuming three archetypes of stations (2, 6, and 10 tons per day), the different rates of utilization became clear. The following key findings emerged:

- *Excluding stations that are expected to recieve less than 1 t/d of demand leads to a large increase of utilization rates while coverage does not drop much*
- *The three archetypes of 2, 6 and 10 t/d are predicted to cover a realistic network of HRS*
- *Key corridors and area's of interest are identified based on synthetic truck traffic flows*

Support for these conclusions can be found in section 12.1.3. The initial distribution of the stations inherently leads to 100% of coverage. Thereafter, both analyses (on utilization and coverage vs. number of stations) pointed toward the exclusion of stations that draw less than 1 ton per day of demand. These locations showed low utilization for more than 70% of 2 t/d HRS. Moreover, with regards to the objective of minizing CAPEX, it allows to build half of the stations while still covering well over 90% of demand. From this final set of locations, DACH, Benelux and the UK emerged as the most interesting areas. Important corridors in Poland, France, Italy and Switzerland should also be mentioned. Portugal, Norway and Finland are among the most uninteresting locations for investment.

What is the influence of hydrogen backbone supply on the optimal locations of European hydrogen refueling stations?

The fifth sub-question aims to include supply in the network design. As introduced, this included determination of the influence but also checking for optimization of the previously determined locations. To get there, costs of supply by backbone were calculated by use of the Optimal Network Layout Tool. This yielded the following key findings:

- *The growing backbone will increasingly be able to cover a large portion of demand from FC-HDTs*
- *Average station sizes and degree of interconnection are main drivers of cost reduction*
- *A low levelized cost of hydrogen transportation can be found in pipelines*

Support for these conclusions can be found in section 12.1.4. The analyses lead to the conclusions that the backbone over time is capable of meeting much of Europe's demand. However, the 2030 and 2040 backbone would need to be supported by other supply methods. By 2050 the backbone is expected to be capable of covering all of central Europe via pipeline (totaling coverage of 65% of expected demand). The output of the Optimal Network Layout Tool provide a visual design of the network system. Based on this network, combined with the capacities and lengths that were specified, average station sizes and degree of interconnection emerged as the main drivers of cost reduction. This key finding leads to the conclusion that HRS location planning should be done with a future network design in mind. The 2030 and 2040 designs show little interconnectivity, as not many stations come online. By upgrading the 2030 station locations in terms of average size, few additional stations need to be built in 2040. This leads to lower levelized transportation costs of hydrogen, as few new pipelines need to be built while more customers can be served. Tactical placement of these stand-alone stations in terms of supply, leaves additional room for large cost reductions due to interconnection in the future. Interconnectivity will increase by 2050, and a well-designed network early on will prove cost-effective over time.

As indicated, an attempt is made to optimize the demand-based network of station locations. This optimization should be based on the cost of supply according to the optimal network design and the cost of demand. The latter is determined by the increase in fuel costs for rerouting after relocation of an HRS. Unfortunately, the optimization of the network system design could not take place. Therefore, the point of supply of each HRS is assumed to be located directly on the backbone. This yielded the following key findings:

- *Costs of supply by distribution pipelines are reasonably balanced with costs of demand*
- *Network optimization in terms of supply has significant effect on hydrogen pricing*
- *Absolute distance increase for FC-HDTs is relatively low*

Support for these conclusions can be found in section 12.1.5. Keep in mind that not the optimal network is optimized, but a more simple "direct connection" network. The findings lead to the conclusion that there is a delicate balance between the costs of supply and demand on which network planning should be based. The cost figures used for this study appear to be balanced, making it possible to optimize towards locations of value. These new locations reanalyzed as a network system design. The optimized locations are found to cause a large decrease in the cost of supply. A drop that is most noticeable in the levelized price for hydrogen, which can be noticed at the pump. On the other hand, the absolute distance for FC-HDTs to reach their HRS does remain within limits. However, the impact of the average distance increase on customer traction is not made clear. The overall take out of these findings is that supply by pipeline should be considered in network planning at an early stage, as it will prove cost effective over time.

What are critical uncertainties to the development of a European hydrogen refueling network?

The final sub-question acts as a call to mark unpredictability in the system. In addition, it acts as a reflection on the modeling. The exploration of scenarios was deemed appropriate for both. Support for these conclusions can be found in chapter 13. Future demand and technological advancements in hydrogen were identified as critical uncertainties. The scenario analysis for demand revealed importance of the market penetration rate on the AFIR implementation (as previously discussed). An incorrect prediction of market penetration could lead to large financial losses due to underutilization. In addition, the best case scenario analysis showed how this station location model responds to uncertainty in demand. If the average size of stations does not increase, a rapidly rising market penetration rate would lead to an almost unfeasible number of stations to be built by 2030.

For the supply, the technological factor of the pipes was tested. Changing the assumption from transmission pipelines to distribution pipelines revealed the sensitivity of the model. An increase in costs of pipes made the costs of supply heavily outweigh costs of demand. It is made clear, that for a balance to be present in costs, only distribution pipelines need to be considered. This entails that no full backbone extenstion must be sought for solely mobility. Only regional extension of the backbone with distribution types yields a balance. This is reaffirmed by the analysis that coverage of the distribution network catches up with the transmission network over time. The 2040 transmission network looks like the 2050 backbone with distribution pipelines added to it. An interesting (but obvious) finding that also emerged was that, when supply is uncertain, a connected backbone adds value.

14.2. Discussion and Suggestions for future research

In section 12.2, multiple assumptions and limitations of the models are discussed. This discussion touches upon the most relevant limitations. In addition to these inherent limitations, there are many uncertainties in the variables used in modeling. This problem arises in any attempt to model real life, but it is even more relevant with hydrogen. Very little accurate (/historical) data is available because this is an infant and highly uncertain market. The scenario analysis of chapter 13 is a first attempt to identify these uncertainties and test the model for outcomes. This immediately reveals the sensitivy of the model towards the input parameters. To place the conclusions in context, and allow more crips recommendations to be made, the most important limitations and uncertain parameters are highlighted below. On top of that, each section immediately hints on future research or other models to address these limitations or uncertainties. The results of the scenario analysis and validation are taken into account.

Model limitations

The validation showed that the database underlying the thesis is quite robust. The shortest-path routing assumption seems to hold ground when regarding long-haul transport. However, the growth factors to predict increasing volumes remain quite insecure. Factors are determined per country, making the data more precise but also more vulnerable. The economic (under)development of one country can cause input-output figures to look completely different in three decades. More disaggregated models (as described in chapter 5) could improve the research. However, disaggregate data on European scale are very difficult to obtain and process.

For the station location modeling, the largest point of discussion surrounds the pre-determination of the number of HRS for k-means modeling. In this study, this was determined by dividing the total demand by the average size of the stations. This yields a realistic total of stations when current available archetypes are known. However, it also creates a high dependence on both the market penetration rate and the average station size (of which the former is highly uncertain). For this study, the validation showed that the number of HRS for 2030 (2 t/d), is reasonably in line with expectations. The scenario analysis showed, how the model responds to an increase in market penetration rate. As a result of this fundamental dependence, the model only represent differences in topology when fluctuating the market penetration. The relative distributions of under- or overutilization and distances to an HRS remain the same. Testing different station size for each decade could provide additional insights. Otherwise, there are many different methods for station location modelling. Flow refueling location models are most popular among these. Use of such a study has its own drawbacks but can provide insight into these factors.

For the network system design, redundancy and a forced balance of demand and supply were identified as relevant limitations. The drive to optimize cost-effectiveness inherently leads to a lack of redundancy. However, redundancy is rather expensive and only when the risks are really great will it be incorporated. Hence, it is assumed redundancy might only be accounted for within the backbone and not in the pipes towards stations. As design of the backbone falls out of scope of this study, the limitation becomes irrelevant. The limitation of requiring supply and demand to be equal at each time step also fall out of scope of this study: supply is assumed to come from one point and be equal to demand at all times. This leaves no relevant limitations of the network system design, as supply points and the backbone design are scoped out. For potential future research regarding supply, they can be taken into account. Realistic supply points would add a lot of value to the network, especially in combination with research on different supply lines.

The main limitation in the two-step optimization lies in the cost balance. The scenario analysis shows that the balance between costs of supply and demand is relatively delicate. If either one of the costs is estimated too high, this will influence results greatly. For this study, a balance was present, but very often unsatisfactory solutions are the result of these types of optimization. This is why the optimizations are generally described by a globalizing or utility function. All the individual distances would be combined into a single global value, which requires a study by itself to be on point.

INPUT PARAMETERS

As mentioned earlier, there are many uncertainties associated with modeling a hydrogen market. The market is volatile, non-existent and price sensitive. On top of that, depending on who you ask, you get different predictions on input parameters. There seems to be a divide between hydrogen enthusiasts and skeptics. Input uncertainties aside, there are certain assumption in the input that still need to be discussed. When calculating the distance from a datapoint (representing trucks) to an HRS, the distance is calculated "as the crow flies". This is briefly mentioned, but no consequences are yet indicated. What this assumption leads to is an underestimation of the distances that trucks must travel to reach an HRS. Costs of demand would see a large increase, and the balance that is discussed will lean more towards demand than it currently does. Since supply is now predominant, this may even produce a better balance. To overcome these limitations, there are many routing APIs that calculate distances based on actual navigation routes. However, they were found to be computationally too heavy for this study. A different data set, or a better model to calculate these data, could make such a study possible.

Another major but undiscussed assumption is the backbone itself. The assumed design of the backbone is based on a thesis, but there are many different designs. The design was incorporated due to its fit with the methodology. This particular design does not extend much beyond southern Europe (into e.g. Spain). This seems logical from a mobility perspective, since there is relatively low demand in those countries. However, the scenario analysis showed that a backbone dedicated solely to mobility is not viable from a cost perspective. Industries and other sectors, which may require further backbone expansion, must be taken into account. The thesis behind the backbone did exactly that, but it is of course based on yet a series of assumptions. Running the model with different backbones as a basis might result in interesting insights in the levelized costs of transportation as well as the topology.

Something that is not addressed at all in this study, but is very relevant, is operational uncertainty. The short-term fluctuation in hydrogen demand, volatility in prices and costs, and utility bills are types of operational uncertainties. Since this study modeled at a very aggregate level, these have not been taken into account. However, not modeling over the time period of a year, but over a period of 7 days will lead to very different demands for station archetypes. Fuel demand fluctuates greatly throughout the day, and phenomena such as the balance between peak capacity and underutilization will play a large role in network planning. The inclusion of diesel refueling behavior data could provide more insight into the impact of this uncertainty.

The final limitation lies in the research objective of the study. The research objective caused this study to focus primarily on where demand was located. After all, the initial goal was to eliminate the biases of Shell diesel data for demand. With that information, an initial network was set up and influence of supply was sought second. However, a more complete study would add more supply-side research. The backbone design that forms the basis of these analyses was also designed with supply sources in mind. These specific sources are not considered in this study. Lines of supply might look very different if they were. The inclusion of supply sources and the resulting options for different supply lines per station will further approximate reality. Supply costs will be more specific and tailored to each station location. In this way, a complete network can be designed, with a backbone for Central Europe, but also supplying HRS in the more peripheral areas by tube trailer.

14.3. RECOMMENDATIONS

This final section attempts to summarize all the research, findings and conclusions into detailed concrete recommendations for Shell.The recommendations are divided into several sections and are presented in the order in which they emerged from the study. For completeness and overview, one relevant figure is given for each recommendation and they are given their own page.

For Shell, it is important that the results are properly incorporated into network planning. In this light, it is recommended that the results are treated as layers in a larger design. The first layers to be added could be realistic new station locations (greenfields) and the current locations of Shell stations that can be retrofitted or expanded. Placing these layers over each other reveals the real possibilities for optimal station locations. Further studies are recommended to determine more layers of the puzzle. The most obvious opportunity seems to lie in the specification and diversification of supply. Specific hydrogen supply sources are not considered and different distribution types may fit each station.

Lobby for subsidies on the peripheral regions of Europe

A small part of this research focused on analysis of the AFIR. However, if negotiations succeed and the AFIR is implemented, it should serve as the basis for all network planning. Therefore, it is of utmost importance to operators like Shell that the regulation does not enforce unfeasible plans. Currently, hydrogen refueling stations are suggested every 150 km along TEN-T corridors and in every urban node. The allocation of projected demand for 2030 (5% market penetration) yielded the following results:

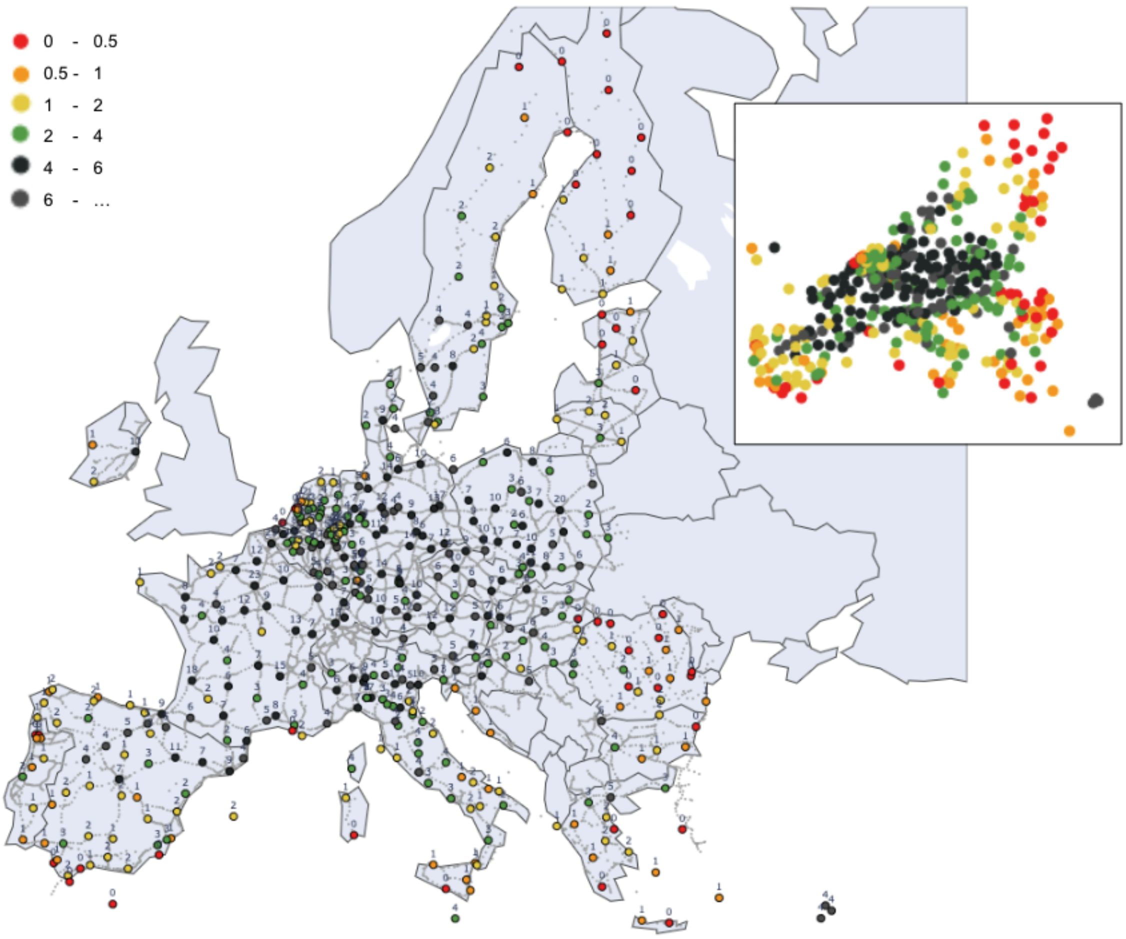


Figure 14.1: AFIR network - allocated demand to show over- and underutilization

Based on these results, the recommendation for Shell is to lobby for subsidies on stations in the peripheral areas of Europe. It is not advisable to lobby strictly against these stations, since not many will be seriously underutilized. A pan-European network with full coverage will help get long-haul fuel cell transport off the ground. Therefore, it may be better to support the regulation and apply for grants in regions where it is needed to be viable. It should be noted that most of the underutilized stations are located in areas where Shell has little or no activity. If competitors were to fill in these gaps, it could help kick off the market while Shell doesn't have to get its hands on it.

If one were to lobby for a change of the AFIR, it is recommended to take a look at the list of urban nodes instead of the distance threshold. The distance threshold seems to add little obligations as the urban nodes cover the lion share of the TEN-T corridors. Furthermore, most stations show signs of overutilization. The two ton per day ask should not be a restriction in any sence for most of Europe.

START OUT WITH DEMAND-BASED LOCATIONS, BUILD THE NETWORK WITH SUPPLY IN FRONT OF MIND

To generate traction for the heavy-duty hydrogen market, it is important to start building the hydrogen refueling network with preferable locations for demand. This provides a somewhat natural design and reduces initial financial losses in the start-up phase. To do so, the 2030 network design of this thesis functions as a solid base:

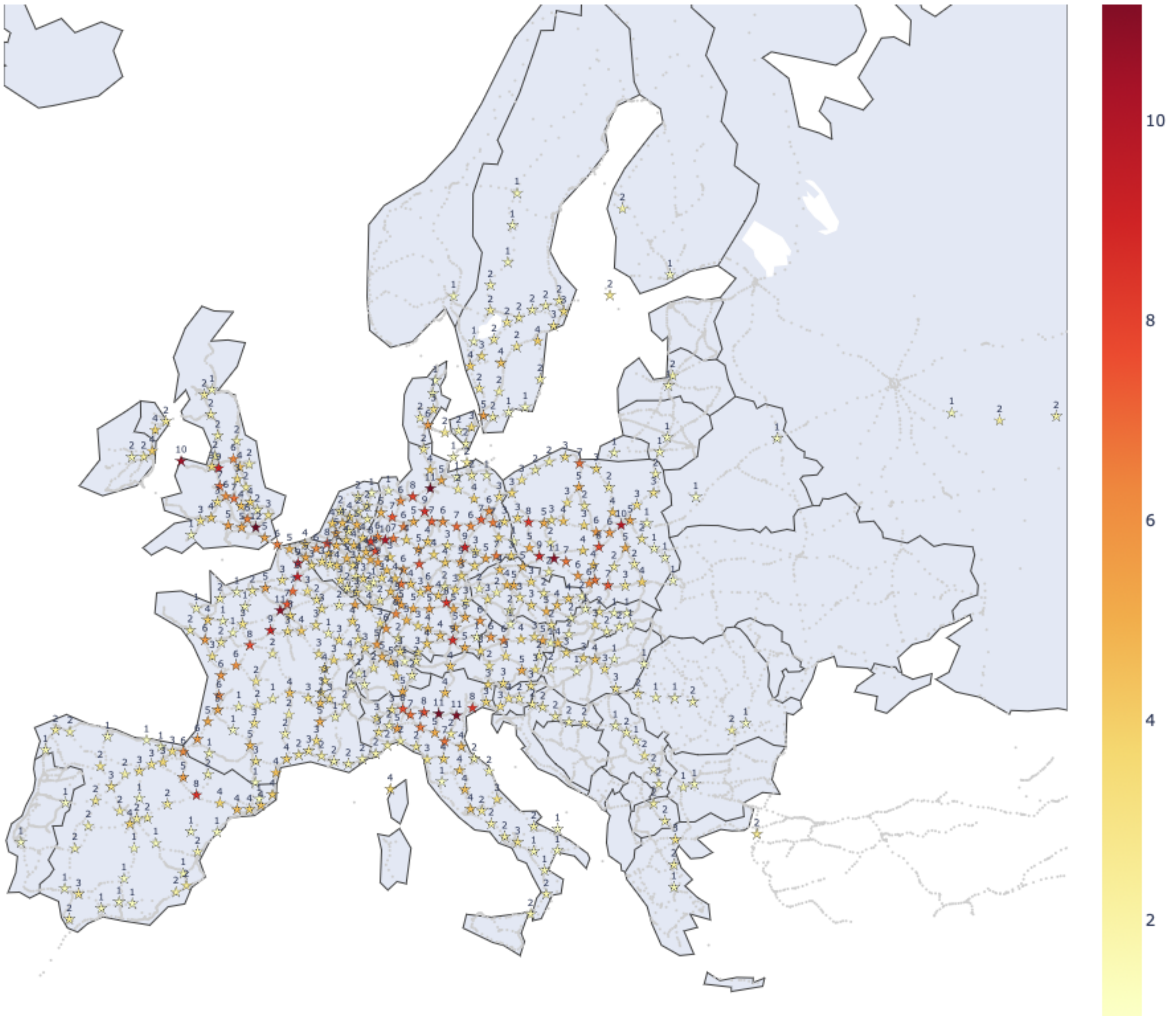


Figure 14.2: Demand-based network - indicative coloring for station size

In terms of station size, archetypes of 2, 6, and 10 tons per day are recommended. Assuming an average station size of 2 tons per day, these sizes seem to create a general coverage network. Stations in locations that fit the archetype of 2 tons per day have the greatest risk of underutilization. However, this is due to the fact that they are designated along less busy roads. Regardless, locations with potential demand for larger archetypes should hold the upperhand. These locations can later on be extended to fit a larger archetype, which can reduce relative costs of supply.

DACH, Benelux and the UK were identified as the most interesting area's for initial investment. High potential also resides along key corridors in Poland, France, Italy and Switzerland. It is advised Shell looks into the creation of a network along these highways that connects central Europe. However, the exact decisions should be made with the future of supply in front of mind as the final recommendation will reveal.

DESIGN AN INTEGRAL NETWORK IN AN EARLY STAGE, BASED ON THE DISTANT FUTURE

Earlier it was recommended that the stations of the startup phase be based on demand. However, proper planning of the network with respect to supply, yields huge cost savings over time. Transport through the backbone will be relatively cheap if there is sufficient demand. It is therefore recommended that the backbone is accounted for prior to its presence. The early phase demand-based stations should be based on the overlap of the 2030 demand based network and the 2050 supply optimized network which is displayed below. To do so, a full-scale network should be designed today, for 2050.

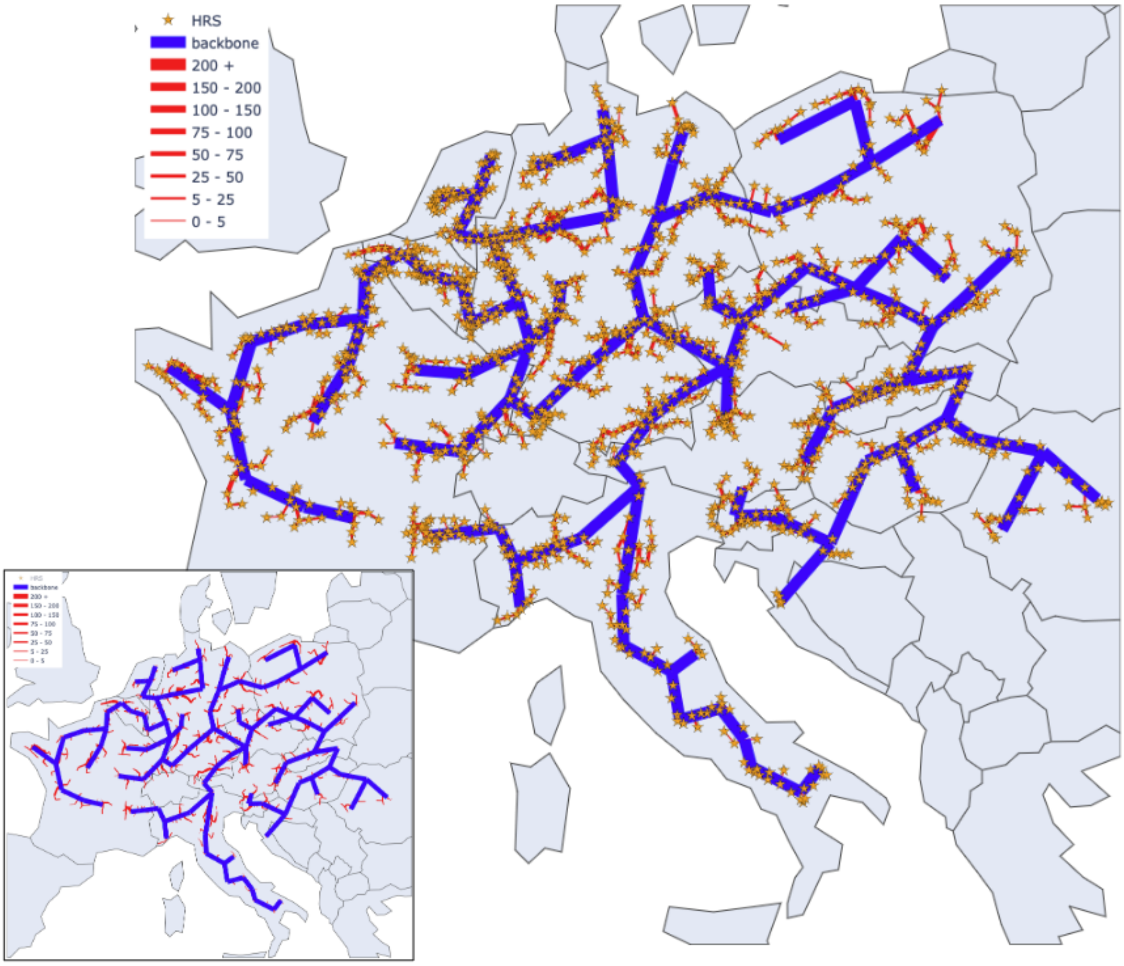


Figure 14.3: Network system design of two-step optimization

Once a backbone and initial stations are in place, the size of the stations and the degree of interconnection appear to be the key factors for further cost reductions. The former can be tapped by growing the stations from 2030 to 2040. To reap the benefits of standardization, it is recommended that this be done by expanding the standard archetype from 2 to 6 t/d. In order to facilitate this growth, the necessary space must be available. Therefore, greenfields should be acquired today to ensure the cost reductions of tomorrow. This study reveals that the levelized cost of hydrogen transportation is greatly reduced in this way. However, growing stations will not increase coverage or minimize driving distances for FC-HDTs. For this reason, it is recommended that the average station size remains the same after 2040.

When demand increases but the average size of stations remains the same, more stations should be built. Once a backbone is in place, the degree of interconnection to the backbone appears to be a key factor for cost optimization of these new stations. A high degree of interconnection leads to relatively low supply costs, since pipeline capacity can be shared. The 2050 network in the figure shows how a proper design allows for capacity sharing of pipelines. Station locations should be determined with this front of mind.

BIBLIOGRAPHY

[1] European Commission, *A clean planet for all: A European long-term strategic vision for a prosperous, modern, competitive and climate neutral economy*, Tech. Rep. (2018).

[2] Hydrogen Council, *Path to hydrogen competitiveness - A cost perspective*, Tech. Rep. (2020).

[3] R. Nugroho, P. K. Rose, T. Gnann, and M. Wei, *Cost of a potential hydrogen-refueling network for heavy-duty vehicles with long-haul application in Germany 2050*, International Journal of Hydrogen Energy **46** (2021), 10.1016/j.ijhydene.2021.08.088.

[4] Hydrogen Council, *Hydrogen Insights - A perspective on hydrogen investment, market development and cost competitiveness*, Tech. Rep. (2021).

[5] Fuell Cells and Hydrogen - Joint Undertaking, *Hydrogen cost and sales prices,* (2022).

[6] Y. Zhou and S. Searle, *Cost of Renewable Hydrogen Produced Onsite at Hydrogen Refueling Stations in Europe,* (2022).

[7] European Commission, *TENtec Interactive Map Viewer,* (2021).

[8] R. Huisman, *Towards a robust European hydrogen network*, Ph.D. thesis, Delft University of Technology (2021).

[9] Eurostat, *NUTS - Nomenclature of territorial units for statistics - background,* (2021).

[10] D. Speth, V. Sauter, P. Plötz, and T. Signer, *Synthetic European road freight transport flow data,* Data in Brief **40**, 107786 (2022).

[11] J. Zhang, L. Liu, Y. Fan, L. Zhuang, T. Zhou, and Z. Piao, *Wireless Channel Propagation Scenarios Identification: A Perspective of Machine Learning,* (2020).

[12] ACEA, *Interactive map – Truck hydrogen refuelling stations needed in Europe by 2025 and 2030, per country,* (2022).

[13] E. Rivard, M. Trudeau, and K. Zaghib, *Hydrogen storage for mobility: A review,* (2019).

[14] P. Kluschke, T. Gnann, P. Plötz, and M. Wietschel, *Market diffusion of alternative fuels and powertrains in heavy-duty vehicles: A literature review,* (2019).

[15] International Energy Agency, *The Future of Hydrogen – Analysis - IEA,* International Energy Agency (2020).

[16] S. Siegemund, M. Trommler, O. Kolb, V. Zinnecker, P. Schmidt, W. Weindorf, W. Zittel, T. Raksha, and J. Zerhusen, *E-Fuels study. The potential of electricity-based fuels for low-emission transport in the EU,* Deutsche Energie-Agentur GmbH (dena) (2017).

[17] NASA, *The Effects of Climate Change,* (2021).

[18] IPCC, *Climate Change 2014 Mitigation of Climate Change*, Tech. Rep. (Cambridge University Press, 2014).

[19] IPCC, *limate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Tech. Rep. (Cambridge University Press, Cambridge, 2013).

[20] European Commission, *Climate change consequences,* (2021).

[21] IMF, *World Economic Outlook*, Tech. Rep. (International Monetary Fund, 2021).

[22] F. Birol, *COP26 climate pledges could help limit global warming to 1.8 °C, but implementing them will be the key*, Tech. Rep. (IEA, 2021).

[23] Eurostat, *Wind and water provide most renewable electricity*, Tech. Rep. (2021).

[24] European Commission, *REPowerEU: A plan to rapidly reduce dependence on Russian fossil fuels and fast forward the green transition**,* (2022).

[25] D. Teichmann, W. Arlt, and P. Wasserscheid, *Liquid Organic Hydrogen Carriers as an efficient vector for the transport and storage of renewable energy,* International Journal of Hydrogen Energy **37** (2012), 10.1016/j.ijhydene.2012.08.066.

[26] K. Walkowicz, A. Meintz, and J. Farrell, *R&D Insights for Extreme Fast Charging of Mediumand Heavy-Duty Vehicles*, Tech. Rep. (National Renewable Energy Laboratory, 2019).

[27] Shell, *Decarbonising road freight: Getting into gear,* (2021).

[28] R. Wurster and P. Schmidt, *Renewable Hydrogen in Fuel Cell Heavy Duty Trucking* , Tech. Rep. (Ludwig-Bölkow-Systemtechnik GmbH (LBST), 2020).

[29] European Clean Hydrogen Alliance, *Mission & Vission,* (2020).

[30] Shell, *Shell starts up Europe's largest PEM green hydrogen electrolyser,* (2021).

[31] Shell, *SHELL HYDROGEN,* (2021).

[32] Shell, *H2Accelerate – new collaboration for zero emission hydrogen trucking at mass-market scale,* (2020).

[33] h2accelerate, *Expectations for the fuel cell truck market,* (2021).

[34] M. A. Nicholas, S. L. Handy, and D. Sperling, *Using geographic information systems to evaluate siting and networks of hydrogen stations,* in *Transportation Research Record*, 1880 (2004).

[35] J. E. Kang and W. Recker, *Strategic hydrogen refueling station locations with scheduling and routing considerations of individual vehicles,* Transportation Science **49** (2015), 10.1287/trsc.2014.0519.

[36] A. Shukla, J. Pekny, and V. Venkatasubramanian, *An optimization framework for cost effective design of refueling station infrastructure for alternative fuel vehicles,* Computers and Chemical Engineering **35** (2011), 10.1016/j.compchemeng.2011.03.018.

[37] Y. Li, F. Cui, and L. Li, *An integrated optimization model for the location of hydrogen refueling stations,* International Journal of Hydrogen Energy **43** (2018), 10.1016/j.ijhydene.2018.08.215.

[38] P. K. Rose and F. Neumann, *Hydrogen refueling station networks for heavy-duty vehicles in future power systems,* Transportation Research Part D: Transport and Environment **83** (2020), 10.1016/j.trd.2020.102358.

[39] S. W. Hwang, S. J. Kweon, and J. A. Ventura, *Locating alternative-fuel refueling stations on a multi-class vehicle transportation network,* European Journal of Operational Research **261** (2017), 10.1016/j.ejor.2017.02.036.

[40] P. Kluschke, R. Nugroho, T. Gnann, P. Plötz, M. Wietschel, and M. Reuter-Oppermann, *Optimal development of alternative fuel station networks considering node capacity restrictions,* Transportation Research Part D: Transport and Environment **78** (2020), 10.1016/j.trd.2019.11.018.

[41] K. Tafakkori, A. Bozorgi-Amiri, and A. Yousefi-Babadi, *Sustainable generalized refueling station location problem under uncertainty,* Sustainable Cities and Society **63** (2020), 10.1016/j.scs.2020.102497.

[42] M. Miralinaghi, B. B. Keskin, Y. Lou, and A. M. Roshandeh, *Capacitated Refueling Station Location Problem with Traffic Deviations Over Multiple Time Periods,* Networks and Spatial Economics **17** (2017), 10.1007/s11067-016-9320-3.

[43] S. Bae, E. Lee, and J. Han, *Multi-period planning of hydrogen supply network for refuelling hydrogen fuel cell vehicles in Urban areas,* Sustainability (Switzerland) **12** (2020), 10.3390/su12104114.

[44] Y. Kuvvetli, *Multi-objective and multi-period hydrogen refueling station location problem,* International Journal of Hydrogen Energy **45** (2020), 10.1016/j.ijhydene.2020.08.126.

[45] S. Lim and M. Kuby, *Heuristic algorithms for siting alternative-fuel stations using the Flow-Refueling Location Model,* European Journal of Operational Research **204** (2010), 10.1016/j.ejor.2009.09.032.

[46] D. Thiel, *A pricing-based location model for deploying a hydrogen fueling station network,* International Journal of Hydrogen Energy **45** (2020), 10.1016/j.ijhydene.2020.06.178.

[47] M. Miralinaghi, Y. Lou, B. B. Keskin, A. Zarrinmehr, and R. Shabanpour, *Refueling station location problem with traffic deviation considering route choice and demand uncertainty,* International Journal of Hydrogen Energy **42** (2017), 10.1016/j.ijhydene.2016.12.137.

[48] J. Hoekman, *Optimizing Hydrogen Refueling Station Locations and Tour Planning: Heavy Weight Road Transport in the Northern Netherlands*, Ph.D. thesis, The University of Groningen (2021).

[49] T. Fleiter, *Climate Protection Scenario 2050 - summary of second final report,* (2016).

[50] P. Gerbert, P. Herhold, J. Burchardt, S. Schönberger, F. Rechenmacher, A. Kirchner, A. Kemmler, and M. Wuensch, *Climate Paths for Germany,* (2018).

[51] O. Bahn, M. Marcy, K. Vaillancourt, and J. P. Waaub, *Electrification of the Canadian road transportation sector: A 2050 outlook with TIMES-Canada,* Energy Policy **62** (2013), 10.1016/j.enpol.2013.07.023.

[52] A. Wang, K. Van Der Leun, D. Peters, and M. Buseman, *European Hydrogen Backbone*, Tech. Rep. (Guidehouse, 2020).

[53] C. Wulf, M. Reuß, T. Grube, P. Zapp, M. Robinius, J. F. Hake, and D. Stolten, *Life Cycle Assessment of hydrogen transport and distribution options,* Journal of Cleaner Production **199** (2018), 10.1016/j.jclepro.2018.07.180.

[54] C. Yang and J. Ogden, *Determining the lowest-cost hydrogen delivery mode,* International Journal of Hydrogen Energy **32** (2007), 10.1016/j.ijhydene.2006.05.009.

[55] J. Zhou, Y. Wu, Y. Tao, J. Gao, Z. Zhong, and C. Xu, *Geographic information big data-driven two-stage optimization model for location decision of hydrogen refueling stations: An empirical study in China,* Energy **225** (2021), 10.1016/j.energy.2021.120330.

[56] M. Adnan Khan, C. Young, and D. Layzell, *The Techno-Economics of Hydrogen Pipelines* , Tech. Rep. (2021).

[57] G. Tezel and R. Hensgens, *HyWay 27: waterstoftransport via het bestaande gasnetwerk?* , Tech. Rep. (Strategy&, 2021).

[58] P. W. Heijnen, A. Ligtvoet, R. M. Stikkelman, and P. M. Herder, *Maximising the Worth of Nascent Networks,* Networks and Spatial Economics **14** (2014), 10.1007/s11067-013-9199-1.

[59] P. W. Heijnen, R. M. Stikkelman, A. Ligtvoet, and P. M. Herder, *Using Gilbert networks to reveal uncertainty in the planning of multi-user infrastructures,* in *2011 International Conference on Networking, Sensing and Control, ICNSC 2011* (2011).

[60] J. Chatzimarkaris, C. Levoyannis, A. Van Wijk, and F. Wouters, *Towards The Creation of the European Hydrogen Economy,* Hydrogen Act (2021).

[61] U. Von der Leyen, *Opening keynote by President von der Leyen at the European Hydrogen Week 2021,* (2021).

[62] D. J. Jovan and G. Dolanc, *Can green hydrogen production be economically viable under current market conditions,* Energies **13** (2020), 10.3390/en13246599.

[63] COM(2019) 640 final, *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS - The European Green Deal,* (2019).

[64] European Council, *Fit for 55,* (2022).

[65] European Commission, *2030 Climate Target Plan,* (2021).

[66] European Commission, *European Climate Law,* (2021).

[67] European Commission, *SWD/2022/230 final - COMMISSION STAFF WORKING DOCUMENT IMPLEMENTING THE REPOWER EU ACTION PLAN: INVESTMENT NEEDS, HYDROGEN ACCELERATOR AND ACHIEVING THE BIO-METHANE TARGETS,* (2022).

[68] European Commission, *Recovery plan for Europe,* (2020).

[69] COM/2020/301 final, *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A hydrogen strategy for a climate-neutral Europe,* (2020).

[70] DIRECTIVE 2014/94/EU, *on the deployment of alternative fuels infrastructure,* OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL (2014).

[71] Regulation (EU) No 1315/2013, *on Union guidelines for the development of the trans-European transport network,* of the European Parliament and of the Council (2013).

[72] COM/2020/824, *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on guidelines for trans-European energy infrastructure and repealing Regulation (EU) No 347/2013,* (2020).

[73] EESC 2020/05835, *Opinion of the European Economic and Social Committee on 'Proposal for a Regulation of the European Parliament and of the Council on guidelines for trans-European energy infrastructure and repealing Regulation (EU) No 347/2013',* (2020).

[74] European Commission, *COM/2021/803 final - Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on common rules for the internal markets in renewable and natural gases and in hydrogen,* (2022).

[75] European Commission, *COM/2021/804 final - Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the internal markets for renewable and natural gases and for hydrogen (recast),* (2022).

[76] Regulation (EU) 2019/1242, *setting CO2 emission performance standards for new heavy-duty vehicles and amending Regulations (EC) No 595/2009 and (EU) 2018/956 of the European Parliament and of the Council and Council Directive 96/53/EC,* of the European Parliament and of the Council (2019).

[77] Regulation (EC) No 595/2009, *on type-approval of motor vehicles and engines with respect to emissions from heavy duty vehicles (Euro VI) and amending Regulation (EC) No 715/2007 and Directive 2007/46/EC and repealing Directives 80/1269/EEC, 2005/55/EC and 2005/78/EC,* of the European Parliament and of the Council (2009).

[78] European Commission, *European Transport Policy. Europe on the Move – Volume 2,* (2019).

[79] Commission Regulation (EU) 2017/2400, *implementing Regulation (EC) No 595/2009 of the European Parliament and of the Council as regards the determination of the CO2 emissions and fuel consumption of heavy-duty vehicles and amending Directive 2007/46/EC of the European Parliament and of the Council and Commission Regulation (EU) No 582/2011,* (2017).

[80] Regulation (EU) 2018/956, *on the monitoring and reporting of CO2 emissions from and fuel consumption of new heavy-duty vehicles,* of the European Parliament and of the Council (2018).

[81] European Commission, *European vehicle emissions standards – Euro 7 for cars, vans, lorries and buses,* (2020).

[82] Directive (EU) 2018/2001, *on the promotion of the use of energy from renewable sources,* of the European Parliament and of the Council (2018).

[83] Directive 2009/30/EC, *amending Directive 98/70/EC as regards the specification of petrol, diesel and gas-oil and introducing a mechanism to monitor and reduce greenhouse gas emissions and amending Council Directive 1999/32/EC as regards the specification of fuel used by inland waterway vessels and repealing Directive 93/12/EEC,* of the European Parliament and of the Council (2009).

[84] Power and Water, *NETWORK PLANNING CRITERIA ,* (2003).

[85] Y. Chen, J. Li, K. Sheng, T. Yang, X. Xu, Z. Han, and Z. Hu, *Many-Criteria Evaluation of Infrastructure Investment Priorities for Distribution Network Planning,* IEEE Access (2020), 10.1109/ACCESS.2020.3043248.

[86] L. Li, H. Manier, and M. A. Manier, *Hydrogen supply chain network design: An optimization-oriented review,* (2019).

[87] J. A. Ventura, S. J. Kweon, S. W. Hwang, M. Tormay, and C. Li, *Energy policy considerations in the design of an alternative-fuel refueling infrastructure to reduce GHG emissions on a transportation network,* Energy Policy **111** (2017), 10.1016/j.enpol.2017.09.035.

[88] Y. Honma and M. Kuby, *Node-based vs. path-based location models for urban hydrogen refueling stations: Comparing convenience and coverage abilities,* International Journal of Hydrogen Energy **44** (2019), 10.1016/j.ijhydene.2019.03.262.

[89] W. Bai and L. Zhang, *How to finance for establishing hydrogen refueling stations in China? An analysis based on Fuzzy AHP and PROMETHEE,* International Journal of Hydrogen Energy **45** (2020), 10.1016/j.ijhydene.2019.12.198.

[90] R. M. Pendyala, V. N. Shanker, and R. G. McCullough, *Freight travel demand modeling: Synthesis of approaches and development of a framework,* Transportation Research Record (2000), 10.3141/1725-02.

[91] W. J. Baumol and H. D. Vinod, *An Inventory Theoretic Model of Freight Transport Demand,* Management Science **16** (1970), 10.1287/mnsc.16.7.413.

[92] B. Bayliss, *The Measurement of Supply and Demand in Freight Transport,* (1988).

[93] Eurostat, *Database,* (2022).

[94] R. H. Lin, Z. Z. Ye, and B. D. Wu, *A review of hydrogen station location models,* International Journal of Hydrogen Energy **45** (2020), 10.1016/j.ijhydene.2019.12.035.

[95] L. A. Meira, F. K. Miyazawa, and L. L. Pedrosa, *Clustering through Continuous Facility Location Problems,* Theoretical Computer Science **657** (2017), 10.1016/j.tcs.2016.10.001.

[96] M. Kuby and S. Lim, *The flow-refueling location problem for alternative-fuel vehicles,* Socio-Economic Planning Sciences **39** (2005), 10.1016/j.seps.2004.03.001.

[97] M. Kuby, L. Lines, R. Schultz, Z. Xie, J. G. Kim, and S. Lim, *Optimization of hydrogen stations in Florida using the Flow-Refueling Location Model,* International Journal of Hydrogen Energy **34** (2009), 10.1016/j.ijhydene.2009.05.050.

[98] Y. W. Wang and C. R. Wang, *Locating passenger vehicle refueling stations,* Transportation Research Part E: Logistics and Transportation Review **46** (2010), 10.1016/j.tre.2009.12.001.

[99] I. Capar, M. Kuby, V. J. Leon, and Y. J. Tsai, *An arc cover-path-cover formulation and strategic analysis of alternative-fuel station locations,* European Journal of Operational Research **227** (2013), 10.1016/j.ejor.2012.11.033.

[100] F. Plastria, *CONTINUOUS LOCATION PROBLEMS : RESEARCH, RESULTS AND QUESTIONS,* (1993).

[101] A. Arora, *K-means for Beginners: How to Build from Scratch in Python,* (2021).

[102] P. W. Heijnen, E. J. Chappin, and P. M. Herder, *A method for designing minimum-cost multisource multisink network layouts,* Systems Engineering **23** (2020), 10.1002/sys.21492.

[103] Q. K. Wan and H. K. Lo, *A Mixed Integer Formulation for Multiple-Route Transit Network Design,* Journal of Mathematical Modelling and Algorithms **2** (2003), 10.1023/B:JMMA.0000020425.99217.cd.

[104] P. Luathep, A. Sumalee, W. H. Lam, Z. C. Li, and H. K. Lo, *Global optimization method for mixed transportation network design problem: A mixed-integer linear programming approach,* Transportation Research Part B: Methodological **45** (2011), 10.1016/j.trb.2011.02.002.

[105] J. Lu, A. Gupte, and Y. Huang, *A mean-risk mixed integer nonlinear program for transportation network protection,* European Journal of Operational Research **265** (2018), 10.1016/j.ejor.2017.07.025.

[106] X. Wang and Q. Meng, *Discrete intermodal freight transportation network design with route choice behavior of intermodal operators,* Transportation Research Part B: Methodological **95** (2017), 10.1016/j.trb.2016.11.001.

[107] N. Johnson and J. Ogden, *A spatially-explicit optimization model for long-term hydrogen pipeline planning,* International Journal of Hydrogen Energy **37** (2012), 10.1016/j.ijhydene.2011.08.109.

[108] C. Gil, R. Baños, J. Ortega, A. L. Márquez, A. Fernández, and M. G. Montoya, *Ant colony optimization for water distribution network design: A comparative study,* in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 6692 LNCS (2011).

[109] A. S. Mohaymany and A. Gholami, *Multimodal feeder network design problem: Ant colony optimization approach,* Journal of Transportation Engineering **136** (2010), 10.1061/(ASCE)TE.1943-5436.0000110.

[110] A. A. El-Saleh, M. Ismail, R. Viknesh, C. C. Mark, and M. L. Chan, *Particle swarm optimization for mobile network design,* IEICE Electronics Express **6** (2009), 10.1587/elex.6.1219.

[111] A. Babazadeh, H. Poorzahedy, and S. Nikoosokhan, *Application of particle swarm optimization to transportation network design problem,* Journal of King Saud University - Science **23** (2011), 10.1016/j.jksus.2011.03.001.

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

[113] Eurostat, *Transport Database,* (2020).

[114] E. Szimba and M. Kraft, *ETISplus Database Content and Methodology,* (2012).

[115] M. Charitonidou, *E-Road Network and Urbanization: A Reinterpretation of the Trans-European Petroleumscape,* Urban, Planning and Transport Research **9** (2021), 10.1080/21650020.2021.1950045.

[116] C. Upchurch and M. Kuby, *Comparing the p-median and flow-refueling models for locating alternative-fuel stations,* Journal of Transport Geography **18** (2010), 10.1016/j.jtrangeo.2010.06.015.

[117] P. Fränti and S. Sieranoja, *How much can k-means be improved by using better initialization and repeats?* Pattern Recognition **93** (2019), 10.1016/j.patcog.2019.04.014.

[118] M. A. Syakur, B. K. Khotimah, E. M. Rochman, and B. D. Satoto, *Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster,* in *IOP Conference Series: Materials Science and Engineering*, Vol. 336 (2018).

[119] M. S. Rahman, *Basic Graph Terminologies,* (2017).

[120] V. N. Evstigneev, V.A.Kasyanov, *Graph Theory for Programmers: Algorithms for Processing Trees,* (1985).

[121] C. F. Bazlamaçci and K. S. Hindi, *Minimum-weight spanning tree algorithms A survey and emphirical study,* Computers and Operations Research **28** (2001), 10.1016/S0305-0548(00)00007-1.

[122] J. B. Kruskal, *On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem (1956),* in *Ideas That Created the Future* (2021).

[123] P. Heijnen, *Optimal Network Layout Tool,* (2022).

[124] P. Kluschke and F. Neumann, *Interaction of a Hydrogen Refueling Station Network for Heavy-Duty Vehicles and the Power System in Germany for 2050,* (2019).

[125] Directive 96/53/EC, *Laying down for certain road vehicles circulating within the Community the maximum authorized dimensions in national and international traffic and the maximum authorized weights in international traffic,* .

[126] European Union, *Heavy goods vehicles (over 3.5 t) - standard speeds limits in Europe,* (2015).

[127] P. Rose, *Modeling a potential hydrogen refueling station network for fuel cell heavy-duty vehicles in Germany in 2050*, Ph.D. thesis, Karlsruher Instituts für Technologie (2020).

[128] C. Cunanan, M. K. Tran, Y. Lee, S. Kwok, V. Leung, and M. Fowler, *A Review of Heavy-Duty Vehicle Powertrain Technologies: Diesel Engine Vehicles, Battery Electric Vehicles, and Hydrogen Fuel Cell Electric Vehicles,* (2021).

[129] J. J. Gangloff, J. Kast, G. Morrison, and J. Marcinkoski, *Design Space Assessment of Hydrogen Storage Onboard Medium and Heavy Duty Fuel Cell Electric Trucks,* Journal of Electrochemical Energy Conversion and Storage **14** (2017), 10.1115/1.4036508.

[130] U.S.Department of Energy, *Hydrogen Storage,* (2022).

[131] M. Hirscher, *Handbook of Hydrogen Storage: New Materials for Future Energy Storage* (2010).

[132] Idealhy, *Liquid Hydrogen Outline,* (2013).

[133] Hyzon Motors, *ZERO EMISSION VEHICLES,* (2022).

[134] Carpos et al., *Energy, transport and GHG emissions - Trends to 2050* (2016).

[135] P. Carlo Dos Reis, *'Clean' hydrogen production: a nascent market and several promising technologies,* (2021).

[136] IEA, *ETP Clean Energy Technology Guide,* (2021).

[137] L. Van Der Burg, *FROM GREY AND BLUE TO GREEN HYDROGEN,* (2021).

[138] European Commission, *OMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS*, Tech. Rep. 9 (2020).

[139] P. Carlo Dos Reis, *Hydrogen demand: several uses but significant uncertainty,* (2021).

[140] Office of ENERGY EFFICIENCY & RENEWABLE ENERGY, *Gaseous Hydrogen Compression,* .

[141] S. Giddey, S. P. Badwal, C. Munnings, and M. Dolan, *Ammonia as a Renewable Energy Transportation Media,* ACS Sustainable Chemistry and Engineering **5** (2017), 10.1021/acssuschemeng.7b02219.

[142] SAE International, *SAE J2601, Rev 2020, Fueling Protocols for Light Duty Gaseous Hydrogen Surface Vehicles,* SAE International (2020).

[143] H. Barthelemy, M. Weber, and F. Barbier, *Hydrogen storage: Recent improvements and industrial perspectives,* International Journal of Hydrogen Energy **42** (2017), 10.1016/j.ijhydene.2016.03.178.

[144] C. Chardonnet, V. Giordano, L. De Vos, and F. Bart, *Study on Early Business Cases for H2 in Energy Storage and More Broadly Power To H2 Applications,* EU Commission (2017).

[145] M. Van Berkum, *Projectplan & Veiligheidsplan H2 tankstation Breda,* (2019).

[146] W. Mars, *Kostencalculaties in het beroepsgoederenvervoer over de weg,* Panteia (2019).

[147] Rijksoverheid, *ADR 2021,* (2021).

[148] R. Van Rossum, J. Jens, G. La Guardia, A. Wang, L. Kuhnen, and M. Overgaag, *A EUROPEAN HYDROGEN INFRASTRUCTURE VISION COVERING 28 COUNTRIES*, Tech. Rep. (2022).

[149] American Gas Association, *How Does the Natural Gas Delivery System Work?* (2022).

[150] E. Moreno Ruiz, L. Valsasina, D. FitzGerald, K. Treyer, G. Bourgault, G. Wernet, F. Brunner, and A. Symeonidis, *Documentation of changes implemented in the ecoinvent database v3.5,* .

[151] B. Wöffen, *Kosten von Wasserstoffpipellines am Beispiel der EuWaK-Wasserstoffpipeline in Bottrop,* (2010).

[152] S. S. Bhogilla and H. Niyas, *Design of a hydrogen compressor for hydrogen fueling stations,* International Journal of Hydrogen Energy **44** (2019), 10.1016/j.ijhydene.2019.02.171.

[153] Z. Tian, H. Lv, W. Zhou, C. Zhang, and P. He, *Review on equipment configuration and operation process optimization of hydrogen refueling station,* (2022).

[154] W. Yu, X. Dianbo, F. Jianmei, and P. Xueyuan, *Research on sealing performance and self-acting valve reliability in high-pressure oil-free hydrogen compressors for hydrogen refueling stations,* in *International Journal of Hydrogen Energy*, Vol. 35 (2010).

[155] M. Verheijen, *Policy analysis for the supply chain business case of hydrogen refueling stations*, Ph.D. thesis, Delft University of Technology (2021).

[156] I. Uehara, *Separation and Purification of Hydrogen,* Encyclopedia of Life Support Systems **I** (2009).

[157] M. Nordio, S. A. Wassie, M. Van Sint Annaland, D. A. Pacheco Tanaka, J. L. Viviente Sole, and F. Gallucci, *Techno-economic evaluation on a hybrid technology for low hydrogen concentration separation and purification from natural gas grid,* International Journal of Hydrogen Energy **46** (2021), 10.1016/j.ijhydene.2020.05.009.

[158] S. Liguori, K. Kian, N. Buggy, B. H. Anzelmo, and J. Wilcox, *Opportunities and Challenges of Low-Carbon Hydrogen via Metallic Membranes,* (2020).

[159] E. H. Majlan, W. R. Wan Daud, S. E. Iyuke, A. B. Mohamad, A. A. H. Kadhum, A. W. Mohammad, M. S. Takriff, and N. Bahaman, *Hydrogen purification using compact pressure swing adsorption system for fuel cell,* International Journal of Hydrogen Energy **34** (2009), 10.1016/j.ijhydene.2008.12.093.

[160] h. budiharjo, S. Siregar, l. mucharam, darmadi, F. Anindhita, and d. supriyatman, *Increasing of Gas Capacity in Pipeline Using Line Packing Technique for Sustaining Gas Supply to Consumers,* (2006).

[161] J. Zheng, X. Liu, P. Xu, P. Liu, Y. Zhao, and J. Yang, *Development of high pressure gaseous hydrogen storage technologies,* International Journal of Hydrogen Energy **37** (2012), 10.1016/j.ijhydene.2011.02.125.

[162] R. Ortiz Cebolla, B. Acosta, N. De Miguel, and P. Moretto, *Effect of precooled inlet gas temperature and mass flow rate on final state of charge during hydrogen vehicle refueling,* International Journal of Hydrogen Energy **40** (2015), 10.1016/j.ijhydene.2015.02.035.

[163] J. Xiao, X. Wang, P. Bénard, and R. Chahine, *Determining hydrogen pre-cooling temperature from refueling parameters,* International Journal of Hydrogen Energy **41** (2016), 10.1016/j.ijhydene.2016.06.084.

[164] F. Piraino, D. Blekhman, M. Dray, and P. Fragiacomo, *Empirically verified analysis of dual pre-cooling system for hydrogen refuelling station,* Renewable Energy **163** (2021), 10.1016/j.renene.2020.10.004.

[165] K. Reddi, A. Elgowainy, and E. Sutherland, *Hydrogen refueling station compression and storage optimization with tube-trailer deliveries,* International Journal of Hydrogen Energy **39** (2014), 10.1016/j.ijhydene.2014.09.099.

[166] COM/2020/789 final, *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Sustainable and Smart Mobility Strategy – putting European transport on track for the future,* (2020).

[167] COM/2020/562 final, *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Stepping up Europe's 2030 climate ambition Investing in a climate-neutral future for the benefit of our people* , (2020).

[168] European Commission, *The Recovery and Resilience Facility,* (2021).

[169] COM/2020/575 final, *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN CENTRAL BANK, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE, THE COMMITTEE OF THE REGIONS AND THE EUROPEAN INVESTMENT BANK Annual Sustainable Growth Strategy 2021,* (2020).

[170] European Commission, *Detailed Assessment of the Member States Implementation Reports on the National Policy Frameworks for the development of the market as regards alternative fuels in the transport sector and the deployment of the relevant infrastructure. Implementation of Art 10 (3) of Directive 2014/94/EU Accompanying the document REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL on the application of Directive 2014/94/EU on the deployment of alternative fuels infrastructure,* (2021).

[171] M. Lambert and S. Schulte, *Contrasting European hydrogen pathways: An analysis of differing approaches in key markets*, Tech. Rep. (The Oxford Institute for Energy Studies, 2021).

[172] F. Simon and T. Dominguez, *HYDROGEN OUTLOOK IN CENTRAL AND EASTERN EUROPE*, Tech. Rep. (Euractiv, 2021).

[173] R. L. Church and R. A. Gerrard, *The multi-level location set covering model,* Geographical Analysis **35** (2003), 10.1111/j.1538-4632.2003.tb01115.x.

[174] O. Karasakal and E. K. Karasakal, *A maximal covering location model in the presence of partial coverage,* Computers and Operations Research **31** (2004), 10.1016/S0305-0548(03)00105-9.

[175] S. Elloumi, M. Labbé, and Y. Pochet, *A new formulation and resolution method for the p-center problem,* INFORMS Journal on Computing **16** (2004), 10.1287/ijoc.1030.0028.

A

System description: fuel cell heavy-duty trucks

This appendix is an addition to section 4.1. It substantiates and constructs the delineation presented in the main text.

A.1. Heavy-duty truck definitions

To start at the basics, it is important to indicate what we mean by heavy-duty trucks (HDT) or heavy-duty vehicles (HDV). It is common practice to classify vehicles based on gross vehicle weight (GVW). However, there does not seem to be an unambiguous definition of HDTs based on GVW. HDVs are divided into vehicles and vehicles with trailers in the EU and China, while in the US they are defined as single vehicles (or trucks). As there does not seem to be standard terminology to go along with, the table of Kluschke and Neumann (2019) is used to set out a definition for HDV for this paper in specific. In this study, the EU vehicle category N3 and (semi) trailer category O4 are considered, since heavy-duty trucks can be considered vehicles of 16 tons and above and only the EU falls within scope. Table A.1 highlights this delineation inside the table of Kluschke and Neumann (2019).

Table A.1: HDV classification

US		EU				China	
Vehicle Category	Weight (t)	Vehicle Category	Weight (t)	Trailers and Semitrailers	Weight (t)	Trucks Weight	Tractors Weight
		N1	<3.5	O1	<0.75		
				O2	0.75 - 3.5		
2b	3.9 - 4.5	N2	3.5 - 12	O3	3.5 - 10	3.5 - 4.5	3.5 - 18
3	4.5 - 6.4					4.5 - 5.5	
4	6.4 - 7.3					5.5 - 7	
5	7.3 - 8.9					7 - 8.5	
6	8.9 - 11.8					8.5 - 10.5	
7	11.8 - 15.0			O4	>10	10.5 - 12.5	
8a	15.0 - 27.2	N3	>12			12.5 - 16	
						16 - 20	18 - 27
8b	>27.2					20 - 25	
						25 - 31	27 - 35
						>31	35 - 40
							40 - 43
							43 - 46
							46 - 49
							>49

A.2. FC-HDT CHARACTERISTICS

Now that the definition of a HDV/HDT for this study has been established, it is time to look at the (potential) characteristics of a fuel cell heavy-duty truck (FC-HDT). As there are very few practical implementations of the FC-HDT and close to no operating trucks in Europe, for the rest of this research, a vehicle is assumed that is designed up to European regulatory standards with currently technological feasible components. Hence, this section provides an overview of the vehicle dimensions, efficiency and energy consumption of the specific FC-HDV considered in this study.

Length and weight

HDTs using European roads must comply with the rules on weights and dimensions laid down in Directive 96/53/EC [125]. This directive is updated in 2019, among others to benefit zero-emission vehicles. It currently lays out that HDVs (bus + trailer) may be 18.75m long, 2.55m wide, and 4.00m high. Furthermore, ANNEX I limits a Two-axle trailer to 18 tonnes, a three-axle trailer to 24 tonnes and vehicle combinations to 40 tonnes. Only the vehicles specified in ANNEX I 2.2.2 c and d are authorized to have a higher limit, but they fall out of the scope of this research. However, recent regulatory changes (EU Directive 2015/719) have allowed for an increase in the maximum weight and length of vehicle combinations that run on alternative fuels or are zero-emission. The allowed increase in length of up to 50 cm causes the maximum length of the considered HDTs to be 19.25 meters. The maximum additional weight is determined by the weight of the alternative fuel or zero emission technology, with a maximum of 1 ton and 2 tons respectively. By regulation, a 'zero-emission vehicle' is defined in point (11) of Article 3 of 2019/1242 as "a heavy-duty vehicle without an internal combustion engine, or with an internal combustion engine that emits less than 1 g CO2/kWh" [76]. FC-HDTs, having no internal combustion engine, are therefore allowed to carry up to 42 tons.

Maximum speed

As for the speed limit, regulations vary from country to country. Since this study deals with international transport, the lowest European maximum speed is taken as the limit. As this research will be focused on a highway network, speed limits for (non-)urban roads and expressways are left out of scope. For type O4 vehicles and trucks over 3.5 tons, the lowest speed limit on freeways in several EU countries is 80 km/h [126].

Hydrogen storage

When looking at the powertrain, the volume of a FC-HDV is almost as large as that of a conventional diesel-HDV and the setup does not differ much from battery electric trucks [127, 128]. However, when taking into account the fuel storage components there is a negative and important difference. The storage of hydrogen on trucks is a major engingeering challenge. Leaving out the fact that the potential design space is very different for each different type of HDV, a major challenge remains [129]. Hydrogen can be physically stored as a gas or as a liquid. The first requires high pressure (for reasons that will be discussed), the second requires extremely low temperatures as the boiling point of hydrogen at one atmosphere pressure is -252.8 °C [130]. Due to the major engineering challenges that remain and potential safety issues, the use of liquid hydrogen as a fuel is left out of scope of this research. Gaseous fuels, unlike liquid fuels, are stored in cylindrical tanks with a uniform pressure distribution that cannot be easily fitted around other vehicle components. Nugroho *et al.* present a CAD model of a potential FC-HDV which will be used as a basis for this thesis:

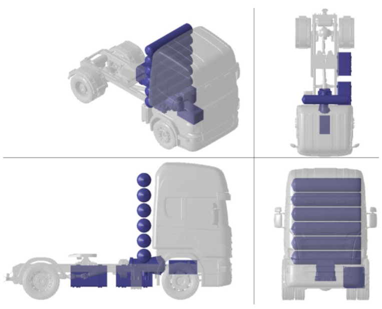


Figure A.1: CAD model of potential FC-HDV [3]

Compressed hydrogen cylinders are bulky and expensive, so to explain why they are still the industry standard, a brief explanation of hydrogens energy density and comparison of storage methods is given. The biggest challenge of hydrogen storage compared to gasoline is inherent to hydrogen: although the specific energy (energy per mass) is much higher than gasoline, the energy density of hydrogen on a volume basis is 4 times lower (8 MJ/l for liquid hydrogen versus 32 MJ/l for gasoline) [130]. Put another way, hydrogen has the highest energy per mass of all fuels, but its low density at ambient temperature leads to low energy per unit volume. Since vehicles are limited in both space and weight, volumetric density becomes an important factor in hydrogen storage design. Therefore, the development of advanced storage methods that allow for higher energy density is required. There are various methods for hydrogen storage for mobility applications:

Method	Gravimetric Energy Density (wt %)	Volumetric Energy Density (MJ/L)	Temperature (K)	Pressure (barg)	Remarks
Compressed	5.7	4.9	293	700	Current industry standard
Liquid	7.5	6.4	20	0	Boil-off constitues major disadvantage
Cold/cryo compressed	5.4	4.0	40–80	300	Boil-off constitues major disadvantage
MOF	4.5	7.2	78	20–100	Attractive densities only at very low temperatures.
Carbon nanostructures	2.0	5.0	298	100	Volumetric density based on powder density of 2.1 g/mL and 2.0 wt % storage capacity.
Metal hydrides	7.6	13.2	260–425	20	Requires thermal management system.
Metal borohydrides	14.9–18.5	9.8–17.6	130	105	Low temperature, high pressure thermal management required
Kubas-type	10.5	23.6	293	120	
LOHC	8.5	7	293	0	Highly endo/exothermal requires processing plant and catalyst. Not suitable for mobility
Chemical	15.5	11.5	298	10	Requires SOFC fuel cell.

Figure A.2: Overview of Storage Methods [13]

The ideal method of storage must enable high volumetric and gravimetric energy densities, rapid fuel uptake and release, operation at room temperature and atmospheric pressure, safe operation, and balanced cost-effectiveness [13]. All current hydrogen storage technologies have significant drawbacks, including complex thermal management systems, boil-off, moderate efficiency, expensive catalysts, stability problems, slow reaction rates, high operating pressures, low energy densities, and risks of violent and uncontrolled spontaneous reactions [13]. As the leading industry standard, compressed hydrogen in cylindrical tanks is much more advanced than the other options. Although it does not match the ideal, it provides a functional solution compared to other technologies. Tank vessels and infrastructure require consideration of material choice, component dimensioning and safety, but resemble established technologies applied to compressed natural gas. Type IV containers of composite materials are commercially available for 350 and 700 bar. Therefore, these will be considered as the standard in the FC-HDT scoped in this thesis.

FC-HDT range

In the same line of reasoning, Nugroho *et al.* (2021) define FC-HDT parameters that comply with European regulations. They determine the storage capability by the available space on an HDV, where EU directive 2015/719 provides an average of 4.3 m3 behind the driver cabin. Their estimate of the loss of space due to use of cylindrical tanks is 50%. This would lead to 2.15 m3 being available for hydrogen storage. A volume of 2.15 m3 hydrogen, under 350 bar considering a gravimetric energy density 23.3 kg/m3 (at 25 C), would be equivalent to 50 kg of hydrogen. Higher pressure, 700 bar, tanks would have a gravimetric energy density of 39.3 km/m3, equivalent to 85 kg of hydrogen in storage [131]. Note that 1 kg of hydrogen contains 33.33 kWh [132]. These amounts of hydrogen would give the FC-HDTs a range of 830 km (350 bar) and 1445 km (700 bar), calculated by assuming a tank-to-wheel (ttw) powertrain efficiency of about 51 % and energy consumption of a fully loaded HDV (2.10 kWh/km). The powertrain efficiency and energy consumption are derived from a comprehensive calculation that is based on multiple sources incorporating weight, efficiency, and volume of the components that make up the truck [127]. These numbers seem to be valid, as one of the few currently available trucks, the Hyzon FCET 8, carries 50-70 kg of hydrogen enabling a range of up to 800km at 350 bar [133]. To the author's knowledge, there are currently no 700 bar trucks available, but a range of 1000+ km is not deemed unfeasible.

A.3. MARKET PENETRATION RATE - LITERATURE REVIEW

As the focus is seldom only on hydrogen, the broader topic of alternative powertrains and fuels in the transport sector is taken as a scope for this literature review. To be consistent with the EU regulatory scheme and to facilitate modeling over time, the timeframes searched for are 2030, 2040, and 2050. When looking at modeling the trucking sector by itself, only two studies appear to focus on the EU and give numbers for FC-HDT penetration for the three timeframes [16, 134].

The EU Reference Scenario of Carpos et al. (2016), projects the impact of macro-economic, fuel price and technology trends and policies on the evolution of the EU energy system, on transport, and on their greenhouse gas (GHG) emissions. To do so, it provides a model-based simulation of a possible future outlook, given the current policy context, based on certain framework conditions, assumptions, and historical trends. For HDVs, it takes into consideration the 2019 Regulation as will be discussed in section 4.4. The results for HDVs show moderate changes for the coming decades. The introduction of CO2 standards and the Clean Vehicles Directive, together with rising fossil fuel prices, cause internal combustion engine (ICE) diesel to drop from 99% in 2020 to 78% in 2050. Their predicted share of electric and fuel cell HDVs is projected to be at around a combined 4% of the total HDV stock by 2050.

Since this scenario is a reference scenario, and thus does not take into account future development and policies, this study will include the study of Siegemund *et al.* (2017) for delineation. In this study, three scenarios based on different assumptions regarding the future development of powertrains and fuel supply are created. The three scenarios are called the Current policy scenario (or Power to Liquid dominated: PtL-dom), the Power-to-gas dominated (PtG-dom) scenario and the Electric-powertrain (eDrives) scenario. The goal of this study, as opposed to the reference study, is to analyse the need for future (renewable) energy to achieve an 80% or 95% reduction in GHG emissions in EU transport by 2050. The Current policy scenario is a conservative BAU scenario based on established fuels, powertrains and infrastructures. It is assumed that internal combustion engines fueled with liquid fuels produced via power-to-liquid technology dominate all transportation modes. In the Power to gas dominated scenario, the focus is on power-to-gas fuels being increasingly used in electrified powertrains. In the most rigid case, the eDrive-dominated scenario, powertrains in all transport modes are electrified. The results, solely for FC-HDTs, are presented in the table below.

Table A.2: Market Penetration Scenarios for FC-HDTs [16]

Truck >12 t	2030	2040	2050
PtL-dom (%)	0	0	0
PtG-dom (%)	5	15	30
eDrives (%)	15	45	80

To validify these numbers, the percentages of market penetration can be compared with the results of other studies on the topic. As stated, the number of studies on the future market penetration of AFP for HDTs are very limited, especially on a specific level as is asked for (EU and fuel cells separately analysed). On top of that, the views on potential market penetration differ. Therefore, this research will place the results of Siegemund *et al.* into the wider literature review of Kluschke *et al.* (2019) to check validity of the results. Kluschke *et al.* (2019) wrote down a summary of market diffusion studies of AFPs in HDVs, including their methods, main findings and policy recommendations. For reasons of comparison, the authors have categorized the scenario results into two clusters: a reference and a climate protection scenario. In the first, small scale implementation of AFPs is predicted, but in the second, AFPs dominate the market. These two clusters are displayed in figure A.3, revealing boxplots of different studies for the share of AFP vehicles in stock over the next decades. The whiskers show the minimum and maximum of all results, while the box contains all values between the quartiles. The solid line represents the median. Both the reference and the climate protection scenarios are consistent on a geographic level, however the authors do note a high degree of uncertainty. To cope with this, they recommend more research into policy measures, and that infrastructure development and energy supply should be included in order to obtain a holistic understanding of modelling AFP market diffusion. All by all, their results give a clear share of AFP in HDV stock, over time, based on many different views.
For the total share of AFPs, Siegemund *et al.* predict 20, 55, and 95% in their "climate protection scenario" (the eDrives scenario) for 2030, 2040, and 2050 respectively. This is within 10% of the median of all studies combined and can thus be seen as valid numbers. Therefore, the share of the AFP share allocated to FC-

HDTs can be considered valid and used as figures for the rest of this study. A note to be made is that the study excluded battery electric vehicles (BEVs) due to range, but it is not the only study to do so [49–51]. Whereas the eDrives scenario could easily be considered wishful thinking, the PtG-dom scenario may be more in line with reality. Therefore, for this study, market penetrations for FC-HDTs of 5%, 15% and 30% are considered for modeling the next decades.

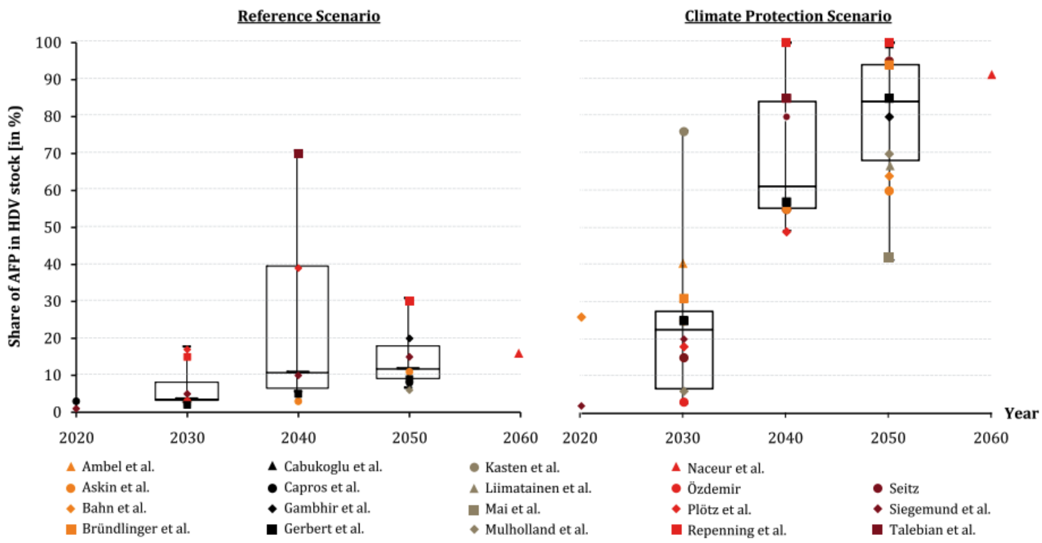


Figure A.3: Market Diffusion of AFP over time in reference and climate protection scenarion [14]

B

System description: supply chain analysis

This appendix is an addition to section 4.2. It substantiates and constructs the delineation presented in the main text.

B.1. Hydrogen Production

First, it is important to shed some light on the production of hydrogen. The molecule by itself is present in very limited quantities. As an atom, however, it is ubiquitous within different molecules. Hydrogen can therefore be produced from these molecules (e.g. water or coal). There are many different technologies for doing this, the best known being electrolysis and gasification. Within these two fields, there are again a lot of different methods: seawater electrolysis, ALK electrolysis, PEM electrolysis, biomass gasification, coal gasificaiton and many more [135]. The main share of these technologies, however, is not commercially available at the moment. The IEA ranks these technologies on a readiness level ranging from 1 to 11, where 1 represents an "initial idea" and 11 represents "proof of stability achieved" [136]. If the technology has a low ranking, it is difficult to formulate credible hypotheses about costs and emissions and its application is out of scope of this research. For each technology, emissions also depend on factors outside the production facility. In this context, the production of the supplied electricity is the most important factor. The production of "clean" hydrogen requires a careful assessment of the available information. Current practice in every day life is to label the cleanliness of hydrogen by color, where three colors are worth mentioning: grey hydrogen is produced by Steam Methane Reforming (SMR); blue hydrogen is produced in the same way but with Carbon capture and storage (CCS), where the carbon dioxide is captured in depleted gas fields or aquifers; green hydrogen is produced by fossil-free electrolysis [137]. It should be noted that emission standards in the industry are expressed in emissions per kg of hydrogen.

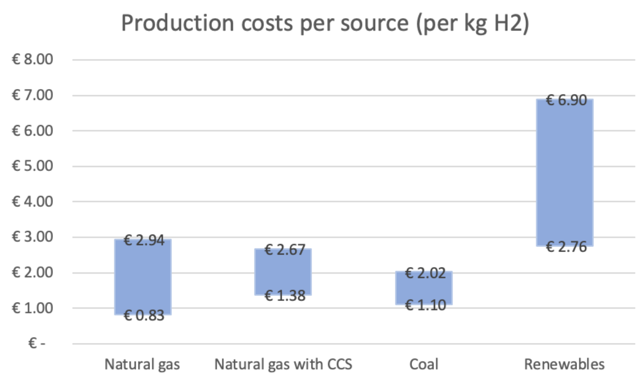


Figure B.1: Production Costs of Hydrogen Today based on International Energy Agency (2020)

Green hydrogen is not yet cost-competitive as opposed to fossil-based grey hydrogen. Figure B.1 shows estimatations of the IEA for production costs per source, translated from USD to EUR. Their estimations, highly dependent on natural gas prices, for grey hydrogen lay around 1.5 €/kg for the EU [138]. If CCS is applied to make the hydrogen blue, costs are estimated around 2 €/kg. Costs for green hydrogen are estimated to lie somewhere between 2.75 - 7 €/kg [15]. When taking into account carbon pricing, a range of € 55-90 per tonne of CO2 would be needed to make fossil-based hydrogen with carbon capture competitive with fossil-based hydrogen today [138]. For purpose of comparison, production costs of blue and grey hydrogen are also displayed in figure B.1

It is expected that costs for renewable hydrogen will go down quickly, as electrolyser costs have already been reduced by 60% in the last ten years, and are expected to halve by 2030. Economies of scale will play a huge role in this effect. Currently, the total production of hydrogen is around 117 Mt (or 3896 TWh) [15]. Less than 0.8 Mt H2 (27 TWh) was produced as green or blue hydrogen, leaving over 99% for the rest. For scaling up the production of clean hydrogen to a 2050 demand (18 to 94 MtH2 (613 to 3115 TWh)), Carlo Dos Reis (2021) identifies six drivers; (1) lower production costs; (2) the availability of sufficient demand for hydrogen; (3) the availability of sufficient production facilities for production; (4) the availability of sufficient 'inputs' – and the relative delivery infrastructure – to produce clean hydrogen from the production facilities deployed; (5) the availability of transport and storage infrastructure linking production facilities to demand sites; (6) policies and regulations acting on the critical drivers identified above, but constrained by environmental sub-factors. These drivers will play a critical role for the infrastructure to become viable, however they are not accounted for in further analysis of this thesis.

DELINEATION

The EU intends to allow only green hydrogen for mobility purposes, and therefore this is the only relevant "color" for this study. Productions cost-estimates between 2.75 and 7 €/kg would fall within scope. As indicated, the cost of production is expected to come down a lot in the coming decades. By how much depends on many factors, and the final values will be of great importance to the future of hydrogen in mobility. However, for station location decisions, as this study was designed to do, these costs are scoped to be not influential. Therefore, further future analysis is not relevant to this thesis.

B.2. HYDROGEN COMPRESSION

Compression of hydrogen plays a role over almost the entire supply chain. Hydrogen is usually produced at relatively low pressure, 20 to 30 bar, and must be compressed prior to transportation [56]. Pipeline transport requires different pressure levels and compressors across the network (>500 km), both depending on the diameter and flow rate required. Typical pressure levels for transportation are further discussed in section B.3. Compressors used today are either positive displacement or centrifugal compressors, with positive displacement compressors being either reciprocating or rotary. These four different type of compressors are shortly touched upon based upon review of the Office of ENERGY EFFICIENCY & RENEWABLE ENERGY [140], followed by a fitting delination for this paper:

- Piston compressors use a motor with a linear drive to move a piston or diaphragm back and forth. This movement simply compresses the hydrogen by reducing the volume it occupies. They are the most commonly used compressor for applications that require a very high compression ratio.
- Rotary compressors compress hydrogen by rotation of gears (or comparable instruments). This type of compression is a challenge for hydrogen because of the tight tolerances required to prevent leakage.
- Ionic compressors act similar to reciprocating compressors but instead of a piston they use ionic liquids. Therefore, they do not require bearings and seals, which are among the most common sources of failure in reciprocating compressors.
- For pipeline applications, centrifugal compressors are preferred as they enable high throughput at a moderate compression ratio. A turbine is rotated at a very high speed to compress the hydrogen. Note that, to compress hydrogen, the speed of the turbine must be three times higher than for natural gas compressors because of the low molecular weight of hydrogen.

DELINEATION

As centrifugal compressors are most suitable for the presumed pipeline, this research will assume this type of compression in its' supply chain. However, since this plays a role before distribution, there is no impact of compression methods on station locations (at this point in the chain). Costs of compression prior to distribution do not have any impact on location choice for the final model.

B.3. HYDROGEN DISTRIBUTION

For this thesis, the distribution is the most relevant step of the supply chain. The method and distance of distribution affect the ideal deployment, hence this section is covered more intensive. So far, only delivery by pipeline has been shortly touched upon. However, the are multiple ways of getting the molecules to the HRS. To provide a comparison this section looks at two distribution methods commonly discussed in the literature for holding the future: supply by pipeline and supply by tube-trailer. On-site production, a third way to get hydrogen to the pump, is excluded from the analysis.The distribution methods included will be technically analyzed and delineated, followed up by an analysis of relevant costs. The latter being built up out of capital expenditures (CAPEX) and operational maintenance expenditures (OPEX). Supply by pipeline is discussed in-depth as relevant techno-economic parameters must emerge.

Joint conditions

As to be able to compare and discuss the different supply chains, joint conditions will be identified. In line with the plans of Shell, this research assumes stations that can provide hydrogen for the bulk of trucks and "regular" private FCEVs. Hence, refueling at 350 and 700 bar must both be technically available at the station. This requires the technical system to store the hydrogen at 880 bar prior to cooling in order to refuel FCEVs with a hydrogen demand of 700 bar [141]. For refueling at 350 bar, 400 bar in storage is required. Furtermore, it is required for both supply chains that the hydrogen is pure enough for application in mobility. Due to the sensitivity of fuel cells, hydrogen should be at least 99.99% pure before being converted into electricity [142]. The background system, such as electricity generation, is assumed to remain constant.

B.3.1. TUBE-TRAILER ROAD DELIVERY

Technical System

When delivered by tube trailer, cylinders are mounted into a trailer, and the HRS is "refuelled" by swapping a trailer of full cylinders with an (almost) empty one. Up until this point in the supply chain, hydrogen is produced at a pressure of 20-30 bar after which it is compressed. Initial storage is likely at a pressure of 200 to 300 bar [143]. Transportation by tube trailer requires the hydrogen to be compressed further to 500 bar to enable supply of 1000 kg per load [144]. It is also possible to deliver hydrogen at lower pressures, which would make one trip cheaper. However, when looking at the volumes delivered, delivery at 500 bar yields a better cost [145]. At the station, the molecules are stored in a cascade made up of multiple pipes.

Analysis of costs

The costs of distributing via tube trailer are variable and determined by the distance of the station to the electrolyser. Note that this distance is dependent on the network of roads, and not "as the crow the flies". One study estimates these costs to be 1.78 €/km [146]. Note that one trip would then be assumed to transport 1000kg of hydrogen, as previously stated. Therefore, to calculate the total costs of supply, the demand of the HRS in questions should be known. To date, hydrogen delivered with tube trailers is often produced with SMR and is therefore not pure enough to be used directly for fuel cells. A Pressure Swing Ad- sorption (PSA) process for the purification step is required. However, since it has already been established that only green hydrogen is within the scope, this is not relevant for this study.

Relevant institutions

Regarding regulations, the same rules apply to the trucks transporting the tube trailer as those discussed for the FC-HDT. The EU guidelines prescribe a limit of 40 tons, which limits the maximum weight of the tube trailer. Furthermore, The European Agreement Concerning the International Carriage of Dangerous Goods by Road (ADR) is in place, as part of United Nations treaty that governs transnational transport of dangerous or hazardous goods [147]. The ADR refers to NEN and ISO standards, which provide a presumption of conformity with the ADR. NEN and ISO standards include (for example) required distances between gaseous storage areas and fire resistance of the storage material.

B.3.2. HYDROGEN BACKBONE

Technical System

At the most fundamental level, pipeline delivery of gas happens through the gas flowing from higher to lower pressure. There are plans to develop a pan-European hydrogen infrastructure (a network of pipelines) which is called the European "hydrogen backbone" [52]. To get there, current gas pipelines can be retrofitted to enable transportation of hydrogen.

Supply of the backbone

The supply of the backbone will have its basis at electrolysers. An analysis on future demand, supply, and transport of hydrogen across Europe [148] estimates sufficient production potential by electrolysis within range of a European backbone. The total expected hydrogen demand could thus be met by green hydrogen produced in the EU and UK and transported over the continent by pipeline. However, they also state that hydrogen imports by pipeline can provide an attractive complement to domestic supply. For the purposes of this study, these production sites were excluded, as this would significantly increase the scope of the study. The backbone itself will be considered the source of supply, without regard to where the backbone gets the hydrogen from.

Hydrogen pipelines

The backbone will be made up out of transmission and distribution pipelines [53]. Hydrogen transmission occurs when hydrogen is transported from a central hydrogen production facility to a single point [54]. Hydrogen distribution occurs from a central hydrogen plant to a distributed network of refueling stations within a city or region. For this study, we can refer to transmission pipelines within the assumed backbone. The end connections from the backbone to the HRS are considered distribution pipelines. The hydrogen energy and capacity values in this section are based on the lower heating value (LHV) as it is customarily in energy system analyses and it enables for straightforward comparison between various fuels.

- Transmission pipelines
  Transmission pipelines typically have a diameter of 6-48 inches and are designed to transport gas over long distances [56]. They often do this at high pressure (10-120 bar), and are generally made out of high-strength steel. The current gas infrastructure consists of pipelines from 20 inch in diameter to 48 inch plus [52]. Converted 36-48 inch pipelines are primarily used for long-distance transportation, and thus are most relevant for the European Backbone. These sizes can transport around 7 to 13 GW of hydrogen respectively (at lower heating value). Regions with smaller average pipe diameters (24-36 inches), including parts of France, Spain and Denmark, will have smaller transportation capacity per pipeline. To make up for lower energy density of hydrogen, the volume flow of hydrogen can be raised compared to natural gas. By doing so, a 48-inch pipeline, can transport around 17 GW in hydrogen and a 36-inch pipeline 9 GW. However, operating at 13 GW for a 48-inch pipeline and 7 GW for a 36-inch pipeline gives much more attractive transport costs per MWh as expensive high capacity compressor stations and corresponding electricity consumption are avoided [52]. These costs increase as compressor stations need be built along each pipeline of significant length (> 100 km). They are there to boost the pressure lost due to the friction of the gas moving through the steel pipe.

- Distribution pipelines
  The pipelines responsible for distribution are typically 2-10 inches in diameter [56]. They are part of a system of pipelines that deliver gas to small industrial plants and customers (like HRS). Within each distribution system, there are sections that operate at different pressures, with regulators controlling the pressure [149] (often at lower pressures: 2-10 bar). Pipes typically range from 2 inches to more than 24 inches in diameter [149] and are consist mainly of plastic [56]. Next to that, no compressors are needed along these pipelines because the destinations are close by. All these traits lead to much lower costs, both CAPEX and OPEX. If the natural gas pipeline distribution is transformed, characteristics of these pipelines can be based on the ecoinvent process for natural gas pipelines [150]. Adaption for hydrogen, would mainly lead to changes in diameter and thickness due to a special coating [53]. This is an alloy that consists of aluminum, zinc and silicon called GALVALUME, which prevents hydrogen diffusion. The lifetime of these pipelines is assumed to be 40 years [151].

The connection between transmission and distribution pipelines often happen at a "gate" station [149]. Here gas is received from many different pipelines and pressure is reduced to the necessary distribution level. On top of that, an odorant might be added to the gas, so leaks can be detected. Lastly, the gate station can also measure gas flow rate to determine the amount being received by the utility [56].

For modeling purposes, it is necessary to know when an HRS is geographically "within the range" of a hydrogen backbone and with which capacity the pipeline should supply. These characteristics determine the length and diameter of the pipeline, which are the largest (non-location-specific) costs [3, 54].

Range of the backbone
Modeling supply by hydrogen pipeline has been done in several forms; Wulf *et al.* (2018) evaluated the environmental impacts of different hydrogen supply chains with differing demand and distances up to 400 km; Zhou *et al.* (2021) considered the transportation costs of hydrogen within 300 km; Yang and Ogden (2007) research transportation costs where transmission distances varied from 25 to 500 km. It is difficult to find a specific value to determine whether an HRS is within "range" of a hydrogen backbone, as a network is modeled. The HRS are possibly interconnected, making a threshold of distance to the backbone itself invaluable. Therefore, the maximum distances from previous research (500 km) is taken to be "within the range" of the backbone.

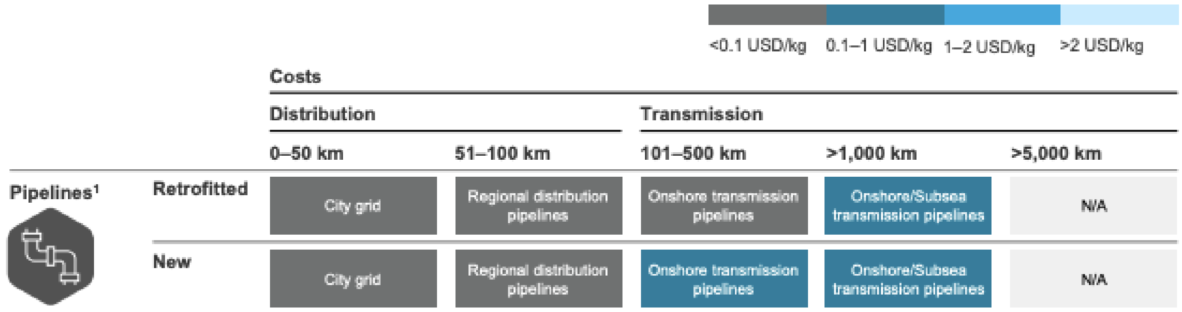


Figure B.2: Pipeline distribution options: distribution prices assuming high utilization [4]

Figure B.2 shows that the 500 km threshold coincides with the aim to keep the levelized cost of hydrogen delivery as low as possible. On top of that, the figure gives threshold values for when to assume pipelines of the distribution or transmission type.

Capacity of the backbone
As already stated, the capacity of the backbone will differ troughout the network along with its diameter. Boundary conditions for both transmission and distribution pipelines have been set. Elaborate calculations for capacity and corresponding costs are given by the technical brief of Adnan Khan *et al.* (2021), but transcend the scope of this study. For modeling purposes, it must be decided whether to consider a transmission or distribution pipeline. The backbone mainly consists of transmission pipelines and will carry the largest part of the transportation. However, this study assumes the presence of a hydrogen backbone and seeks to connect to it. It is assumed that the last 0 to 50 and 50 to 100 kilometers will be transported by distribution pipelines (figure B.2). Hence, distribution pipelines are taken as the standard.

Analysis of costs
As for many large infrastructure networks, high investment costs account for the largest share of the total costs and inflict a barrier for creation. The total installed capital cost of a pipeline include not only materials for the pipeline, but installation costs, rights of way and miscellaneous costs [54].

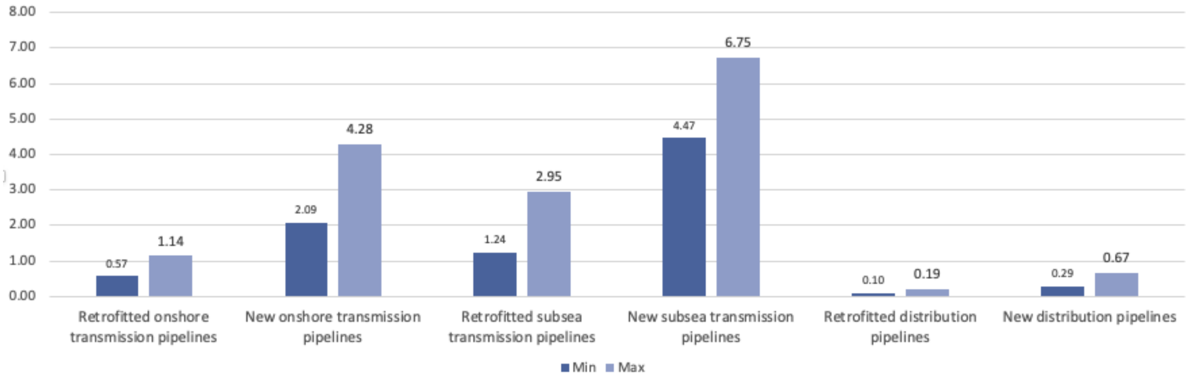


Figure B.3: Overview of different pipeline costs in €M (based on [4])

Installation costs, rights of way and miscellaneous costs fluctuate greatly by region over Europe. However, for modeling purposes they are assumed to be equal. Adnan Khan *et al.* (2021) breaks down these costs and provides formulas to determine each. Their total costs estimates resemble the figures given in figure B.3, hence it assumed all these elements are accounted for. The hydrogen backbone is defined as consisting of pipelines, compressor stations (if present), control valves, and gas metering stations. The figures also reside with the estimates of roughly €2.75 million for new pipelines of Wang *et al.* (2020). There is also validation for the cost of retrofitting, as Tezel and Hensgens 2021 estimate investment costs being reduced to €0.84 million per kilometer. Distribution pipelines are significantly cheaper than transmission pipelines (about 15%), given their smaller diameter and lower pressure.

The costs of retrofitting pipelines depend on a variety of factors including diameter and pressure, the quality of the materials used, the pipeline's overall condition, the existence of cracks, the social costs of construction, and other considerations [4]. Figure B.4 breaks down the possibilities for the different type of pipelines:

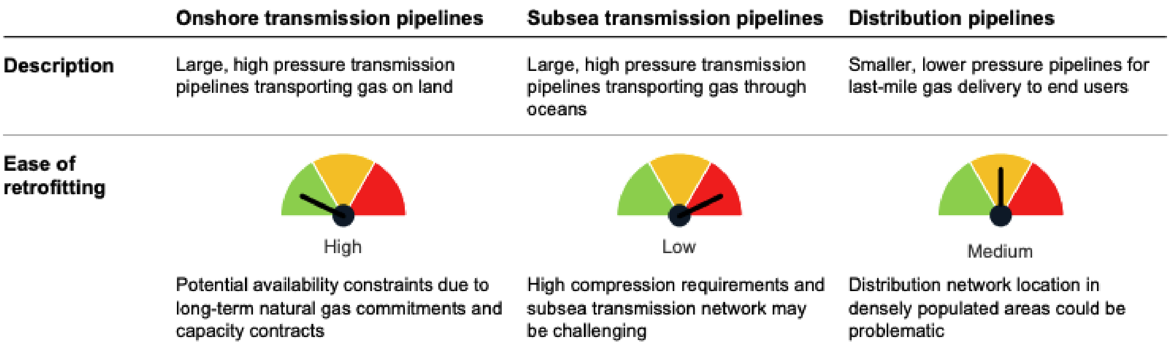


Figure B.4: Pipeline distribution options: possibilities for retrofitting [4]

To calculate the cost of a hydrogen pipeline network, a cost figure per kilomter of pipeline is needed, as well as a figure for capacity expansion. Since distribution pipelines will be the subject of study, these figures are considered. The maximum values are taken as the starting point, as this excludes hydrogen-promoting bias. Since the retrofitting opportunities are ranked as medium (due to densely populated areas), the price per kilometer is taken as the average of the two numbers: € 0.48 M per km. The costs of capacity expansion are less straight forward. As previously discussed, complicated figures for costs of capacity are left out of scope. Therefore, a more simple exponential capacity factor that is used for natural gas infrastructure is considered: 0.6 [58, 59]. This beta () represents a realistic value to indicates the cost advantages to building highcapacity pipelines. It results in the effect that a connection with a capacity twice as much as another connection of the same length is not twice as expensive.

Levelised costs of hydrogen

The levelised costs of hydrogen transportation are estimated to be between €0.09-0.17 per kg of hydrogen per 1000 km, which indicates the method to be cost-effective over long distances across Europe [52]. This is determined using a conversion factor of 1 kg equalling 0.033 MWh (at lower heating value). However, this requires a certain level demand. Adnan Khan *et al.* (2021) established the rule of the thumb that “A demand of 1-1.2 tH2/day/km pipeline is needed to drive economic viability”. When looking at cost optimisation for the total network, the concept of compression versus pipeline dimension and availability of the existing gas network are viewed as the main levers.

Relevant institutions
A pan-European hydrogen backbone will need to be backed up by politics in the form of regulations, incentives and subsidies. The implications will be large, and require holistic management: a review of the gas legislation, policy making on sustainable finance will all play a role in enabling this key European infrastructure. Currently, the European Commission has already announced integration of hydrogen infrastructure planning in their Hydrogen Strategy which will be discussed in section 4.4.2.

DELINEATION
As previously identified, there is a gap in literature surrounding studies of hydrogen refueling station locations supplied by pipeline. Therefore, this is considered the sole source of supply in the system of this thesis. This entails the inclusion of retrofitted gas infrastructure with their dimensions (transmission pipelines) and 98% pure hydrogen. Added to this are the yet-to-be-discussed addition of linepacking for storage, and the resulting need for a PSA for purification. The hydrogen backbone is defined as consisting of transmission pipelines, compressor stations, control valves, and gas metering stations. Hydrogen storage is not taken into account.

For this thesis, the focus lies on the connection of the backbone to the HRS, which is done by distribution piping. A study-based hydrogen backbone will be assumed which will serve as a basis for the model of this study. Hence, the characteristics of the backbone itself do not influence our outcome. Nonetheless, the scope of this study will be determined by the hydrogen backbone considered: it will set limits to the geographical span of the network, over time. Distribution pipelines cost relatively little, at €0.48 M per km. The costs of capacity expansion are determined by exponential capacity factor that is used for natural gas infrastructure: 0.6. Lastly, a discharge pressure of 10-20 bar is entailed. It must be noted that these fundamental network design considerations, are based on analysis of existing infrastructure as well as initial experimental and computational feasibility studies. The resulting generic characteristics do not fully mirror reality as costs will undoubtedly vary across Europe. However, Wang *et al.* (2020) state that, by relying on an infrastructure-based approach and adopting a generic network design for the analysis, the resulting parameters and cost margins are considered representative of the EU average.

With regards to the hydrogen that will be flowing through the system of pipelines (or hydrogen backbone), the EC reported that ultimately it will be all green hydrogen. However, in the short term, there will be a role for low carbon hydrogen [52].

B.4. COMPRESSION AT THE HRS
Compression prior and during transportation has already been discussed. At the HRS site, compression also plays a major role. In fact, it makes up the largest part of the total cost [152]. As the discharge pressure of the hydrogen backbone is assumed to be 10-20 bar, the largest increase in pressure of the supply chain will take place at the HRS. This is due to the necessary dispension of 350 or 700 bar from the HRS into the tank of the FC-HDV. The list of compressors as discussed in section B.2 still constitute the largest share of categorization [153]. For this high-pressure hydrogen compression, currently, piston compressors have the most mature technology and the lowest cost, making it the best choice [154].

Analysis of costs
The necessary compression largely adds to the total expenditures of the HRS. Verheijen ([155]) reports on two studies that indicate a CAPEX of €5000, or 6500, per kg of hydrogen per hour. The OPEX is calculated as 4% of the relative CAPEX. One of the studies does indicate that costs will fall by almost half in the future. However, as for the sections that will follow, these costs are on-site and will not influence the location decision. For this

reason, these costs will not be taken in consideration during modeling and the analysis of costs are no longer discussed.

DELINEATION

As discussed in the FC-HDT characteristics, onboard hydrogen must be stored at 350 or 700 bar to store enough energy to achieve diesel range parity. For on-board gas storage at these pressure levels, the working pressure of the hydrogen storage vessel for 350 bar is 450 bar and the working pressure for 700 bar is 880 bar [153].

B.5. PURIFICATION

On the production side, hydrogen purities from current technologies range from 97.5-98.5% for methane reforming to >99.999% for alkaline and proton exchange membrane (PEM) electrolysis [52]. It is not yet certain what the purity of the hydrogen will be when exiting the pipeline, but it is assumed that it will be of a lower quality than required for use in a fuel cell: 98% [52, 57]. Ultimately, a common specification must be defined for hydrogen transport in Europe, otherwise pipelines will not be interoperable. Nonetheless, as result of transportation of the pipeline, this quality will be of lower purity than the necessary 99.99% for use in fuel cells. Therefore purification must take place at the HRS.

DELINEATION

There are multiple techniques for purification, but only those that fit the hydrogen from the pipeline (±98% purity) will be taken into account [156]. This leaves Pressure Swing Adsorption (PSA) or the use of polymer or metal membranes. The latter both require expensive substances to function [157, 158]. PSA systems are able to produce high purity hydrogen flows at a continuous rate, at the same pressure for output and input flow. On top of that, they have a relatively fast process cycle which fits the purpose for an HRS [159]. A choice for PSA also fits into Shell's strategy and is therefore considered as purification step for this thesis. The purification step, making use of PSA systems, would add €0.36 per kg of hydrogen to the costs in this step of the supply chain [155].

B.6. STORAGE AT THE HRS

Besides distribution, pipelines can also be used to store gas before (and during) transportation. The amount of gas in the pipe is called the "linepack" and the technique of short-term gas storage is called line packing [160]. By raising and lowering the pressure on any pipeline segment, this part of the network can store gas during periods when there is less demand at the end of the pipeline [149]. Hence, linpacking allows handling hourly fluctuations in demand very efficiently. Redundancy of hydrogen supply to consumers would thus see an increase due to higher capability to meet peak demand. The retrofitted pipeline requires that the pressure not fluctuate too much, as that might lead to failure in the material of the pipes [155]. Pressure swings are not to be expected for the demand from HRS, and therefore left out of scope of this research.

The flow to the station and short gas storage in the form of line-packing have been discussed, but before being dispensed onto the vehicles, hydrogen is stored on-site of the refueling station. As discussed, hydrogen can arrive at the HRS in different forms. All types of storage tanks, however, must follow ASME standards. Tube trailer delivery will lead to a trailer of cylinders being present, where liquid will lead to above-ground tanks, which are refuelled by a tanker truck. While tube trailer delivery makes up current practice, this thesis is focused on the use of pipelines and therefore linepacking combined with pressure buffering is considered.

DELINEATION

For an HRS being supplied by a pipeline, the hydrogen storage facility mainly has the function of pressure buffering. The higher working pressure and greater difference in charging pressure will lead to shorter refueling times of FC-HDTs. However, the increase in working pressure will also lead to higher costs, as the compressor will be used more. In accordance with ASME standards, the main part is often made of high-strength structural steel 4130X, and the nominal volume of a single gas cylinder is 0.895 m3 [161].

B.7. PRE-COOLING

Before entering the FC-HDT storage tank, hydrogen must be pre-cooled. The high end state of the charge, to achieve diesel parity filling time, requires low temperatures during the filling process. These temperatures are limited by the safety standards of the storage tank material [162]. On top of that, the protocol for hydrogen dispension (SAE J2601) requires hydrogen to be pre-cooled to the temperature range of 33 to 40 C [142]. In this manner, the temperature onboard of the HDV will stay below its maximum of 85 C during rapid filling [163].

DELINEATION

For the system analyzed, a typical pre-cooler is sufficient. The components present in such system are a hydrogen chiller, a heat exchanger and pipeline system [164]. From the storage vessel, the hydrogen first flows through the heat exchanger in the pipeline before going to the dispenser. Simultaneously, the hydrogen chiller deploys a coolant that enters the heat exchanger, causing sufficient heat exchange. The coolant returns to the cooler, is re-cooled and re-enters the pipeline creating a cycle.

B.8. HYDROGEN DISPENSION

The final part of the supply chain consists of the dispension of the hydrogen into the tank of the HDT. As stated many times, the stations will dispense hydrogen as a compressed gas at pressures of 700 bar and 350 bar depending on the vehicle. When the dispenser is activated, the hydrogen will flow from the storage tanks to the dispenser and then flow through the nozzle into a closed system in the HDT. Before fueling starts, initial safety checks will be performed internally. During this process, the mass, temperature, and pressure inside the vehicle's tank during the filling operation should be monitored, as described in the SAE J2601 refueling protocol [165]. Additional integrity checks will be conducted during periodic breaks during refueling.

DELINEATION

To fill at the speed required to reach diesel parity (10 to 15 minutes), the equipment of the hydrogen filling system must include highpressure filling pipelines, pneumatic shut-off valves and its supporting explosion-proof solenoid valves, electronic pressure regulators, gas guns, temperature and pressure sensors, mass flow meter, monitor, sequential gas control panel and chip controller [153]. These components, mostly integrated in the dispenser, enable the system to protect against overpressure and hose break, compensate to ambient temperate and ensure that the fueling process can be cut off during emergy shutdowns.

B.9. RESULTING SCOPE

The main components of an HRS are; a hydrogen storage system to meet daily demand (in this case, the connection to a hydrogen backbone); a high-pressure buffer storage system; a compressor that pressurizes hydrogen from the storage source pressure to the buffer storage pressure; a refrigeration system that pre-cools the hydrogen before it is dispensed into the HDT tank; and finally a dispenser that manages the flow of hydrogen to the tank. These are all accompanied by the necessary safety equipment and gauges. The figure below shows a schematic representation of the HRS components, when being supplied by a pipeline.

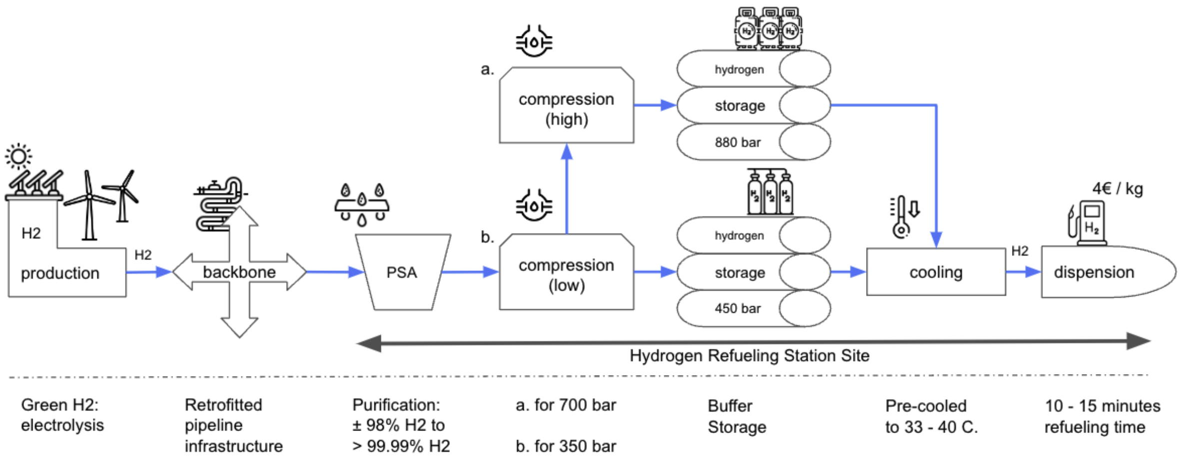


Figure B.5: Schematic Representation of HRS components and supply

The influence on the results for this thesis come from the techno economic parameters of the pipelines. Therefore, these are further visualized to make the scope clear. The distribution pipelines are scoped as the pipes to be constructed. The investment costs for new transmission pipelines would be considered if the assumed backbone were to be expanded.

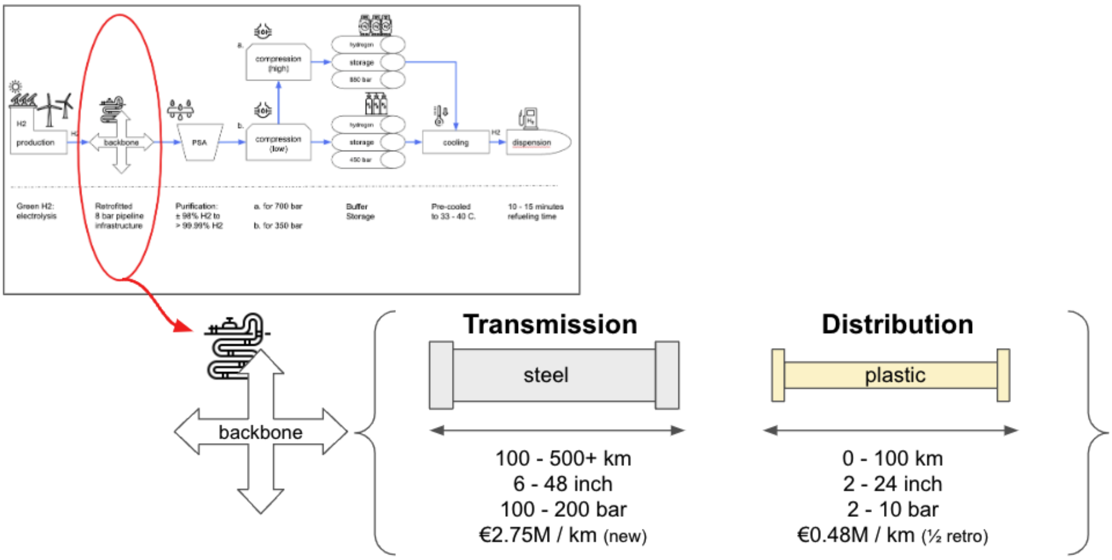


Figure B.6: Schematic Representation of pipeline supply

C

System description: policy analysis

This appendix is an addition to section 4.4. It provides more details on the EU strategies mentioned that were not used for specific purposes in the main study.

C.1. EU strategies

C.1.1. Sustainable and Smart Mobility strategy

In December 2020, the EC presented the Sustainable and Smart Mobility Strategy to "put European transport on track for the future" [166]. The strategy aims to improve the efficiency, interconnectivity and multimodality of the transport system. Within this strategy, the aim is to establish redundant refueling infrastructure for zero-emission vehicles, along with the provision of renewable and low-carbon fuels. It sets out that, achieving these goals will require additional investment of 130 billion euros per year in the transport sector from 2021-2030 [167]. The TEN-T core network alone will require 300 billion over the next decade.

C.1.2. NextGenerationEU

In response to the COVID19-crisis, the NextGenerationEU plans aim to help the EU recover stronger and more resilient [68]. At the core of this plans lay the Recovery and Resilience Facility [168]. Being a temporary recovery instrument, it allows the Commission to raise funds to help Member States implement reforms and investments. It makes available €723.8 billion in loans (€385.8 billion) and grants (€338 billion) for that purpose. These investments should however, be in line with the EU's priorities. The Annual Sustainable Growth Strategy (ASGS) outlines the economic and employment policy priorities for the EU for 12 to 18 months [169]. The "flagships" of the ASGS 2021 (2022 has not been found), address common challenges for EU member states. One of these being the "Recharge and Refuel" pillar. It steers to promote future-proof clean technologies, including refuelling stations, to reach the Paris climate objectives. More specifically, the flagship aims to build 500 hydrogen stations by 2025. The EC strongly encourages member states to include this pillar in their recovery and resilience plans.

C.2. National strategies

Since the directives leave room for member states to implement the objectives through their own strategies, a brief dive into what is currently being planned seems appropriate. The main differences in national policies will therefore be discussed in this section.

C.2.1. AFID implementation

Implementation of the AFID in the member states is checked and progress reported by the EC [170]. National implementation reports show the importance of the AFID, lack of national coordination and low initial investements were identified. Most countries have not met the required targets in 2020, and their strategies differ in level of ambition. Western European Member States have mainly begun started to develop specific strategies for hydrogen HDVs (The Netherlands, Portugal, Germany, France, Spain, Italy).

C.2.2. HYDROGEN STRATEGIES

In the strategies presented based on AFID, the role of hydrogen for HDVs is present in several forms [171]. The Netherlands is aiming to build 50 HRS by 2025, Germany is aiming for 400 HRS by then to add up to 1000 by 2030. France is looking to build 100 by 2023 and 400-1000 by 2028, Spain for 100-150 by 2030 and Italy wants to have 2.5% of its trucks running on fuel cells by 2030. EU cross-border transport is also part of the strategies. Regarding Eastern Europe, several member states have announced hydrogen strategies or transport roadmaps [172].

D

Details on Station Location Models

This appendix is an addition to section 6.2. It provides more detail on the mentioned station location models as distinguished by Lin *et al.* (2020.

D.0.1. The set covering model

For the supply of alternative fuels, most published studies address the problem by maximum covering or set covering approaches. The classical set covering problem aims to find the smallest number of facilities and their locations so that each demand point is covered by at least one facility [173].

D.0.2. The maximal covering location model

The maximal covering location model, as its name suggests, aims to maximize the coverage of a predefined number of refueling stations. It does so by addressing the issue of locating the specific number of facilities such that they maximize the number of demand points that are covered [174]. This entails that demand and traffic flow patterns are not explicitly considered [47].

D.0.3. The p-median model

The p-median model reduces distance to demand points with demand considered as weight (a min-sum problem). It does so, by locating p facilities and allocating demand nodes to them to minimize total weighted distance traveled [116]. It can be used in combination with (e.g.) GIS data to simulate demand [34].

D.0.4. The p-center model

The p-center model, aims to reduce the maximum distance between a node and the station that covers it (min-max) [175]. This causes the model to not be directly linked to demand calculation and therefore not often used for refueling stations but more likely applied to the location of (e.g.) emergency facilities.

D.0.5. The flow-intercepting model.

Flow-intercepting models are focussed on finding the number of refueling stations that maximize the refueled traffic flow between origin-destination (OD) pairs, considering a limited driving range of vehicles. The demand is associated with the traffic flow, which can be calculated in a dynamic way, making it very good for application to refuelling stations. More specifically, the flow-refueling location model (FRLM) [96] is a popular method, as it does consider tanking more than once. The FRLM is path-based, and locates p stations to maximize the number of vehicles on their shortest paths that can be refueled. For a path to be considered refuelable, one or more stations must be located on the path in such a way that the round trip can be completed without running out of fuel, considering the driving range of the vehicle.

E

Station Location Model criteria ranking by Shell

This Appendix sets out the ranking of criteria for the ability of the station location model by Shell. Note that this is different than the performance measures, as it searches for criteria and the modeling ability (e.g. freedom of positions of stations).

E.1. Ranking the criteria

In discussion with Shell, four new criteria are drawn up to rank the models on their applicability for this thesis. These criteria are discussed in the order in which they were ranked:

1. To enable customer traction, **minimizing the driving distance** of the trucks towards the HRS is labeled as the most important trait of a model and thus should be flagged as the objective function. Moreover, this is essential in finding the optimal location based on supply and demand.

2. The freedom a model provides in positioning the stations is ranked second and is aimed to be: **any point**. This is strictly due to the fact that this research functions as an exercise to find locations based on supply and demand. Influence of discrete locations, even in the form of a road network, would interfere with this goal. While it may seem counterintuitive to steer away from highways, coverage may be more optimal in some other way. Proper modeling should serve the answer.

3. The criterium ranked third is the way a model deals with the capacity per station. These could be either capacitated or variable. In reality, however, stations would be **semi-capacitated** as they consist of different archetypes. A capacitated station that has a demand higher than its capacity can be made into a larger archetype. However, this is only up to a certain value (the largest archetype) which should not be exceeded.

4. The number of stations is ranked last. For the models discussed, this can be minimized or fixed (maximization would result in an unrealistic number of stations). Minimizing the number of sites could potentially drive down costs. A model that entails a fixed number of stations would not have this advantage of cost reduction. However, it might be a better reflection of reality as minimization would assume optimal usage of the stations. In reality, stations will be opened up at a certain location and customers will fill up its capacity over time. As both approaches have their relative benefits to the research, it is deemed inessential and the criterium is ranked last.

These findings are summarized as a tablein the main text. The objective functions are highlighted yellow.

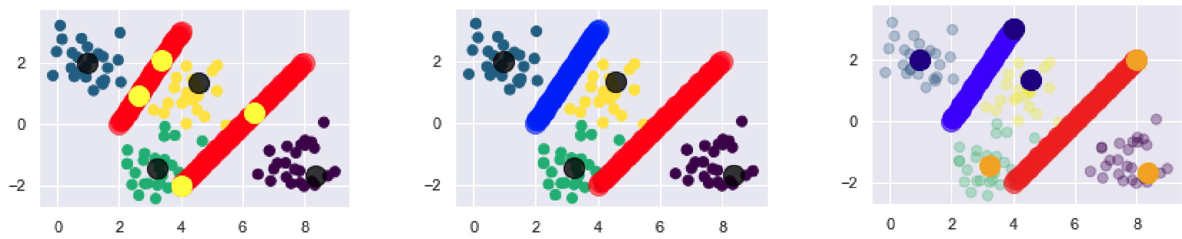
F

EQUALING DEMAND AND SUPPLY

The used Optimal Network Layout Tool needs demand and supply to be equal at all time, the supply source must meet demand exactly. This Appendix explains how this study dealt with doing so for the modeling of the backbone.

F.1. EXPLANATION OF THE PROCESS

Each source that supplies a loose part of the backbone needs to equal exactly what the demand nodes ask from the part of the backbone. In other words, the supply node has to equal the demand nodes that are connected. This must be determined prior to optimization. Figure F.1a shows the same simple example as previously used for k-means explanation but with two pieces of a "backbone added to it in red. It displays how for each centroid the closest point on a part of a "backbone" is found (the yellow point).



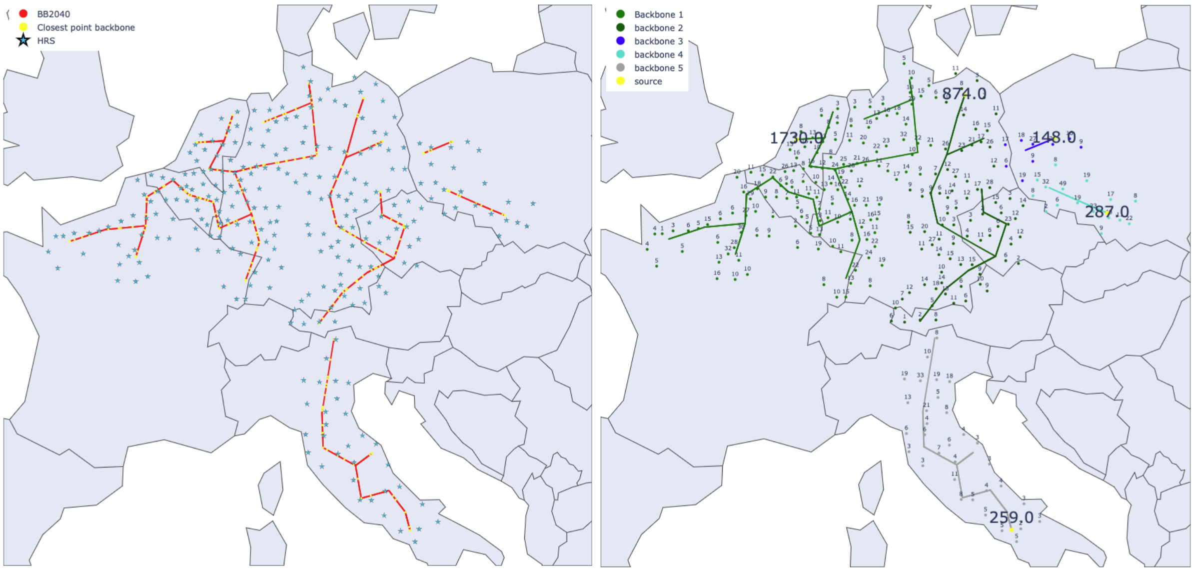
(a) Determine closest backbone

(b) Define separate backbones

(c) Addition of source/supply nodes

Figure F.1: Simple example to define supply per backbone

For each centroid, all of the backbones are checked, which is why they are both colored red. Next, it must be known which piece of the backbone each centroid is connected to. Figure F.1b displays the separated parts of the backbone. The centroids that are closest the each backbone are gathered in a set. The demand of each centroid (HRS) for each backbone must then be summed and allocated as supply to a source node. Figure F.1c colors the centroids the same color as their corresponding backbone and a corresponding supply node is added to the end of each backbone. The figures below show how this was incorporated into the dataset using the data of 2040.



(a) Determine closest backbone 2040

(b) Separated backbones 2040 with respective demand

Figure F.2: Method to define supply backbone 2040

Figure F.2b displays the numbers of demand each separate supply point has to deliver to equal demand of all connected stations. Note that this calculation must be performed prior to usage of the Optimal Network Layout Tool. If the demand and supply of each loose part of the backbone do not match exactly, the tool will model the connection between the parts of the backbone itself.

G

COST FUNCTION OPTIMAL NETWORK LAYOUT TOOL

This Appendix describes the cost function used in the Optimal Network Layout Tool as introduced in section 12.1.4. The function is given below, together with the exact description as it is given by Heijnen [123].

$$C(G) = \sum_{e \in E_n(G)} l_e q_e^{\beta} + spc \cdot s(G) + cpc \sum_{e \in E_o(G)} l_e (q_e - rq_e)^{\beta},$$

Figure G.1: Cost function Optimal Network Layout Tool

Where $E_n(G)$ is the set of all new edges (connections) in the network G, $l_e$is the absolute length (Euclidean distance) of the edge e,q_e is the capacity of this edge and is the capacity-cost-exponent, which indicates how profitable it is to join edges. If capacity does not count in the network costs, = 0, if connections of double capacity are twice as expensive = 1. Normally is somewhere around 0.6. Extra cost can be added for the number of splitting points s(G) in the network. If connections can only be split at existing nodes, spc should be set to a high value.

Also extra cost are calculated for extension of the current (remaining) capacity rq_e of existing connections $E_O(G)$. If cpc = 1 then extending the current capacity of an existing connection is just as expensive as building a new connection. If cpc < 1, extension of the existing connection is cheaper. If extension is impossible, cpc should be set to a high value.

To make the total cost more realistic, you need to multiply it with a reality factor.

Remember that the final minimal-cost-network found by the model might be sub-optimal, since finding the optimal network is hard and we use heuristics to end up with a near-optimal network.