Doctoral theses at NTNU, 2023:313

Šárka Štádlerová

Multi-period facility location problems with capacity expansion: Locating hydrogen production in Norway

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Economics and Management
Dept. of Industrial Economics and Technology
Management

NTNU
Norwegian University of
Science and Technology

Šárka Štádlerová

Multi-period facility location problems with capacity expansion: Locating hydrogen production in Norway

Thesis for the Degree of Philosophiae Doctor

Trondheim, October 2023

Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

ISBN 978-82-326-7326-1 (printed ver.)
ISBN 978-82-326-7325-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:313

Printed by NTNU Grafisk senter

Abstract

According to the Paris Agreement on climate change, CO_2 emissions must be decreased by 40% towards 2030. Decarbonization of the transportation sector is a crucial step to meet the emission reduction targets in the Paris Agreement. Hydrogen is considered a promising low-emission fuel alternative in the transportation sector. However, low experience with hydrogen and limited availability are the main challenges when introducing hydrogen as a low-emission fuel in the transportation sector. Hence, designing the hydrogen production infrastructure and providing a reliable hydrogen supply is essential to encourage potential customers to switch from fossil fuels to hydrogen.

This thesis formulates optimization models and develops efficient solution methods to solve the problem of designing hydrogen production infrastructure in Norway. The papers in this thesis contribute to the state-of-the-art on deterministic and stochastic multi-period facility location problems with capacity expansion. To accurately represent the costs of hydrogen production, we consider specific aspects of hydrogen production technologies in our mathematical models. These aspects include economies of scale in investment, capacity-dependent short-term production costs, and considerations of minimum production requirements. Further, we present solution methods for deterministic as well as stochastic formulations based on Lagrangian relaxation. Furthermore, we provide managerial insight into the cost analysis of hydrogen production in Norway.

This thesis consists of four research papers. In Paper I, we formulate and solve the problem of locating hydrogen production facilities in Norway considering deterministic demand. In Paper II, we present a solution method based on Lagrangian relaxation for the problem introduced in Paper I. In the last two papers, future demand is considered an uncertain parameter. In Paper III, we use sample average approximation to solve the problem, while in Paper IV, a solution method based on Lagrangian relaxation is developed.

Acknowledgements

During the last three years as a PhD student at the Department of Industrial Economics and Technology Management at the Norwegian University of Science and Technology, I have had the pleasure to meet numerous people that influenced my PhD studies.

Firstly, I would like to express my sincere thank to my supervisor Associate Professor Peter Schütz for allowing me to do a PhD under his supervision support. I am grateful for the guidance and time you have spent providing excellent feedback.

My gratitude also goes to my co-supervisor Professor Asgeir Tomasgard for his valuable advice and support.

I would like to thank Associate Professor Sanjay Dominik Jena for hosting me at Cirrelt in Montreal and providing me with valuable and inspiring insides into facility location and capacity expansion problems.

I wish to thank the Research Council of Norway for granting funding through the MoZEES project. Further, my gratitude goes to the Department of Industrial Economics and Technology Management for providing excellent working conditions during my time as a PhD. I am also grateful to our administrative and IT support staff, whose exceptional support has eased my life.

I would like to thank my PhD fellow students and colleagues for creating a friendly working environment, in particular, Vibeke Petersen and my office mate Brede Sørøy.

Finally, I want to express my heartfelt gratitude to thank my family. Their unconditional love and support have been a constant source of motivation and inspiration, especially during challenging times. A special thanks belongs to my boyfriend František Kolovský.

Contents

Introduction

This thesis is motivated by the real-world problem of decarbonising the transportation sector in Norway to meet the CO_2 emission reduction targets. A promising solution to reach these goals is to use hydrogen as an energy carrier in the transportation sector. As of today, hydrogen is predominantly used in industrial processes and produced on-site without an open hydrogen market. Due to the limited demand for hydrogen as a low-emission energy carrier, hydrogen infrastructure other than for industrial processes is not available. Therefore, establishing a reliable hydrogen supply chain is essential to support the hydrogen transition and encourage hydrogen demand from the transportation sector. This thesis addresses the problem of locating hydrogen production facilities in Norway to satisfy domestic hydrogen demand in the transportation sector. This problem can be formulated as a multi-period facility location and capacity expansion problem.

The research in this thesis was performed within MoZEES, a Norwegian Center for Environment-friendly Energy Research (FME), co-sponsored by the Research Council of Norway (project number 257653) and 40 partners from research, industry and the public sector.

Ocean Hyway Cluster provided detailed hydrogen demand data for high-speed passenger ferries, car ferries and coastal route Bergen-Kirkenes, as well as data related to the transition potential of the offshore supply vessels.

This thesis contributes to modelling and solving multi-period facility location and capacity expansion problems. It also provides managerial inside and analysis of hydrogen infrastructure in Norway. The uncertainty in future hydrogen demand further motivates the formulation of a two-stage multi-period facility location and capacity expansion problem for locating hydrogen production. Due to difficulties arising when introducing uncertain demand, we propose solution methods contributing to the literature on solving two-stage stochastic multi-period facility location and capacity expansion problems.

The thesis consists of two main parts. The first part is the introduction. Section 1.1 briefly describes the emission reduction goals in the transportation sector as well as the main steps of the hydrogen supply chain. In Section 1.2, we formulate the research questions and objectives of this thesis. The modelling approach and motivation are summarised in Section 1.3. In Section 1.4, we review related literature and place our papers in context with the state-of-the-art and highlight our contributions to the literature. A summary of the papers included in the thesis is presented in Section 1.5. The second part of this thesis contains the papers themselves.

1.1 Decarbonisation of the transportation sector

Reducing emissions in the transportation sector is a critical and necessary step towards achieving the emission reduction targets outlined in the Paris Agreement on climate change. In 2019, the transportation sector represented about 29% of the global energy demand, almost exclusively covered by fossil fuels. Alternative zero-emission fuels have to replace fossil fuels to achieve the emission reduction targets (DNV GL, 2020). One promising solution is the use of hydrogen fuel cells, which can help to decarbonise the transport sector and significantly reduce greenhouse gas (GHG) emissions, as suggested by Fridstrøm et al. (2018). IEA (2022a) further states that the global energy crisis caused by Russia's invasion of Ukraine accelerates the urgency to use hydrogen, as it contributes to emission reduction targets as well as energy stability.

In 2015, the Norwegian Parliament decided that CO_2 emissions must be decreased by at least 40% (compared to 1990) towards 2030 in an attempt to reach the targets of the Paris Agreement. In 2020, Norway announced the Norwegian climate action plan with an ambitious plan to cut the CO_2 emission by at least 50% (Samferdselsdepartementet, 2021). In 2017, the transport sector in Norway was responsible for emitting 15.8 millions tons CO_2, accounting for 23% of all CO_2 emissions (Aarskog et al., 2020). CO_2 emissions from domestic inland water and coastal transport in Norway alone accounted for 8.7% of emissions from the transport sector in 2018. Therefore, introducing zero-emission energy carriers such as batteries and hydrogen in the maritime and land-based transportation sector can help to reduce CO_2 emissions considerably.

1.1.1 Hydrogen demand

Low experience with hydrogen and uncertain availability of hydrogen fuel may discourage potential customers from switching to hydrogen. Besides, due to a non-existing open hydrogen market, there is no supply chain available. As of today, hydrogen is typically produced on-site for industrial processes. However, to support the hydrogen transition, a reliable hydrogen supply must be ensured. Therefore, it is crucial to adopt policies to initiate the hydrogen transition on the national level. Stimulating the demand side can incline supply chain investments. Further, scaling up hydrogen production and establishing a hydrogen market can enhance the competitiveness of hydrogen in terms of price and availability (IEA, 2022b).

In the maritime sector, the International Maritime Organization presented ambitions to decrease the CO_2 by 50% from 2008 to 2050. Hydrogen and ammonia are the most promising low-emission energy carriers for this sector, as batteries have limited potential for longer distances due to their low energy density (DNV GL, 2020). According to DNV GL (2020), hydrogen has considerable potential in the land-based sector for heavy vehicles travelling long distances as it enables faster refuelling and a longer distance range compared to batteries. For the passenger car fleet, hydrogen is not competitive with batteries.

Norway, like many other countries, is facing a lack of hydrogen infrastructure. Requirements on hydrogen fuels in public transport can generate initial hydrogen demand and encourage the building of hydrogen facilities for reliable local supply (Mäkitie et al., 2020). A suitable candidate to create the initial hydrogen demand in Norway is the sector of high-speed passenger ferries and car ferries that are operated based on public contracts. When contracts are renewed, hydrogen fuel can be required by authorities. Further, high-speed passenger ferries and car ferries operate based on fixed schedules. Therefore the demand has a deterministic character and is easily predictable (Ocean Hyway Cluster, 2020). In the Norwegian maritime sector, hydrogen is considered to be a better alternative than batteries since ferries cross long distances with only a few stops on islands (Ocean Hyway Cluster, 2020). Many islands also do not have a sufficient grid to provide a charging station for ferries. Even if batteries can be suitable for a few ferry connections, which cross only short distances and coming regularly back to the main port, hydrogen is a more convenient energy carrier for most ferry connections (Danebergs and Aarskog, 2020). Introducing hydrogen in maritime transportation can considerably reduce CO_2 emissions and contribute to establishing the hydrogen infrastructure as well as decreasing hydrogen costs. This will further have a positive impact on the implementation of hydrogen in land-based transportation and other sectors (Fridstrøm et al., 2018).

According to DNV GL (2019), hydrogen has considerable potential in the offshore sector and the land-based sector for heavy trucks and long-distance buses in Norway. Further, hydrogen is considered a zero-emission solution for not electrified train connections. Hydrogen demand in the sectors as the land-based transportation sector and the offshore sector is highly uncertain. However, it is expected to increase in the following years (Ocean Hyway Cluster, 2020; DNV GL, 2019). In the land-based transportation sector, the main competing energy carrier are batteries, while ammonia is considered an alternative in the offshore sector. The future energy mix is uncertain, and hence also the hydrogen demand. Therefore, scalability and flexibility are the key features of hydrogen infrastructure to satisfy customers when demand increases (DNV GL, 2019).

1.1.2 Hydrogen production

There are different types of technology to produce hydrogen. To differentiate between the types of hydrogen, colour codes were introduced. Grey hydrogen is produced using steam methane reforming leading to substantial CO_2 emissions as a byproduct and therefore, grey hydrogen is not suitable for decarbonisation purposes. The production process for grey and blue hydrogen is identical. However, blue hydrogen production is complemented by carbon capture and storage, which makes it a nearly carbon-neutral energy carrier (IRENA, 2020). Green hydrogen is produced through water electrolysis using renewable electricity. To mark produced hydrogen as green, it must be produced exclusively from renewable sources without emitting CO_2 emissions in the production process. Figure 1.1 illustrates hydrogen colour classification based on the production process.

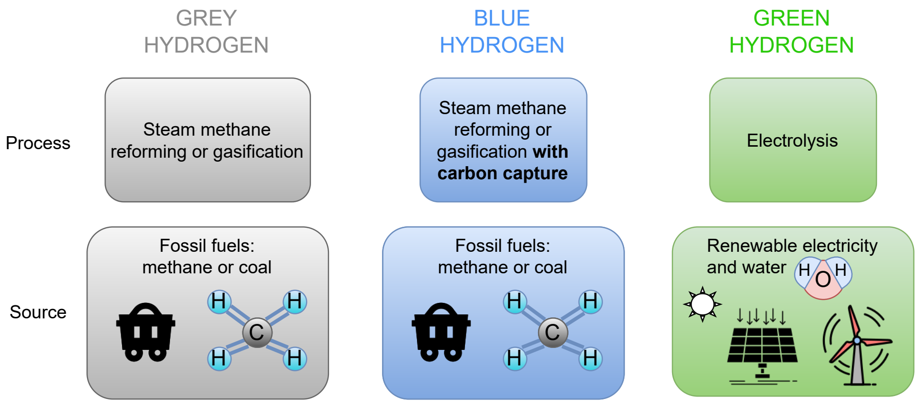


Figure 1.1: Hydrogen colour classification

There are two relevant technologies to produce hydrogen for decarbonisation purposes: electrolysis (EL) and steam methane reforming with carbon capture (SMR+) (Hirth et al., 2019). The current carbon capture technology can capture about 90% of the emitted CO_2, and thus, electrolysis is the only technology capable to produce green hydrogen.

Both SMR+ and electrolysis benefit from economies of scale. Economies of scale arise from scaling up the production quantities and sharing the fixed costs over more units of the produced product. As a result, higher production quantities lead to lower unit costs (Haldi and Whitcomb, 1967). Economies of scale are more pronounced for SMR+. Compared to electrolysis, investment costs in SMR+ are typically higher due to the high costs of the steam methane reforming facility and the carbon capture and storage system. On the other hand, SMR+ is characterised by having lower operational costs than EL. Economies of scale, together with high investment costs make SMR+ a more suitable technology for large-scale hydrogen production, while electrolysis is beneficial for both large and small-scale production (Keipi et al., 2018). Note that the operational costs of SMR+ are to a large degree affected by the price of natural gas, although in the case of EL, the electricity price is the decisive factor (Jakobsen and Åtland, 2016). Alkaline electrolysers have technological limitations associated with minimum production quantities (NEL Hydrogen, 2015). In the future, proton exchange membrane (PEM) electrolysers will provide a higher operational range and faster start-up times than alkaline electrolysers. However, as of today, alkaline electrolysers are the most mature and cheapest production alternative (Andrenacci et al., 2022).

1.1.3 Hydrogen supply chain

Considering the high investment costs related to building a SMR+ facility and economies of scale in production, this technology is rather suitable for a centralised supply chain with a few large strategically positioned facilities and following distribution to customers. Electrolysis can benefit from relatively low investment costs and is therefore profitable for a decentralised supply chain with many small facilities located close to customers enabling on-site refuelling or low distribution costs (Keipi et al., 2018). Figure 1.2 exemplifies the hydrogen supply chain considering blue and green hydrogen. Hydrogen is either transported by trucks to customers or purchased directly on-site in a local facility.

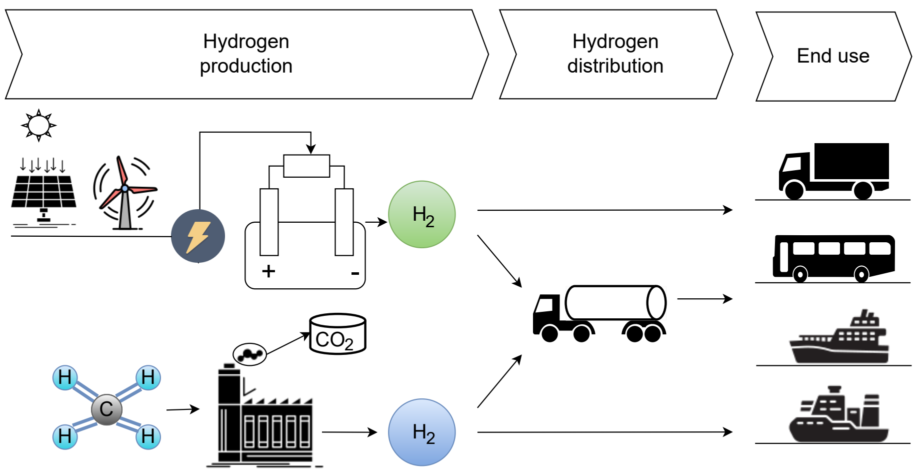


Figure 1.2: Hydrogen supply chain

As of today, the most common method of hydrogen distribution is in the form of gaseous hydrogen transported in high-pressure trucks at a pressure of 300 bar in a 40-foot container (Danebergs and Aarskog, 2020). For the distribution of large quantities, liquid hydrogen is more suitable than gaseous hydrogen due to its higher volumetric density. However, the liquefaction process and the necessity to keep the temperature at −253°C cause high costs. Therefore liquid hydrogen is, as of today, not competitive with other fuels, not even with compressed hydrogen. Liquid hydrogen may enter the market in the medium term together with large-scale hydrogen production taking advantage of economies of scale and advanced liquefaction techniques. Pipelines for hydrogen distribution are considered only for large distances and larger quantities due to high set-up costs. In the medium term, hydrogen can be distributed by ships in gaseous as well as in liquid form (IEA, 2022a).

1.2 Research questions

This thesis formulates optimisation models for locating hydrogen production and proposes efficient solution methods. In this section, we introduce the research questions addressed in the thesis. Resulting of the necessity to start developing the hydrogen infrastructure to encourage potential customers to switch to hydrogen, we formulate our main research questions as follows: Where to locate hydrogen production facilities in Norway? Which capacity and technology to install in? When should investments be made to meet hydrogen demand? Based on the main questions, follow-up problems are formulated.

One of the crucial aspects of this study is how to model and solve the problem of locating hydrogen production. We formulate a mixed-integer facility location and capacity expansion problem. Due to the complexity of the mixed-integer facility location and capacity expansion problem, an efficient solution method is required to solve the problem for larger instances within a reasonable run time.

Considering the uncertainty in future demand and zero-emission requirements on hydrogen production, the resulting question is how to model and solve the problem of designing hydrogen production infrastructure under uncertain demand so that all customers are satisfied in all time periods and all scenarios. Specifically, this includes determining the optimal location, timing of investments, and capacity installation in the first stage and capacity expansion decisions in the second stage. As a result, we formulate a two-stage stochastic multi-period facility location and capacity expansion problem. The difficulty to solve this problem increases compared to the deterministic problem formulation. Therefore, implementing a solution method capable of providing high-quality solutions for large instances within a reasonable run time considering a sufficient number of scenarios becomes necessary.

Further, the properties of hydrogen production motivate the analysis of how to model the non-linear costs in investment and production. Specifically, this includes modelling of economies of scale in investment as well as capacity-dependent non-linear short-term production costs.

The modelling approach and methodology necessary to address the questions above are discussed in Section 1.3.

1.3 Methodology

The decision support method applied in this thesis is mathematical programming, specifically mixed-integer linear programming (MILP). The problems are formulated as a minimising (maximising) function of decision variables subject to equality and inequality constraints. Considering the MILP, some of the decision variables are discrete or binary allowing for modelling a wide range of real-world problems. In general, MILP problems belong to the category of NP-hard problems, and therefore larger instances are hard to solve (Nemhauser and Wolsey, 1999).

In planning processes, some decisions have to be taken without the exact knowledge of future prices or demand for commodities. To address the impact of uncertainty in the decision-making process, two-stage and multi-stage stochastic models with recourse have been developed. These models divide the decision-making process into stages, with new information becoming available at each stage. Recourse actions can be taken accordingly at each stage when uncertainty is disclosed (Birge and Louveaux, 2011; Kall and Wallace, 1994).

This thesis focuses on modelling and solving facility location problems with capacity expansion, considering the impact of economies of scale when expanding a facility. The problem is studied from both deterministic and stochastic perspectives using MILP formulations. A general introduction to static facility location problems and the motivation for formulating multi-period facility location and capacity expansion problems are provided in Section 1.3.1 and Section 1.3.2, respectively. Finally, the impact of uncertainty on the decision process is addressed in Section 1.3.3.

1.3.1 Static facility location

The purpose of facility location problems is to find a trade-off between investment costs in facilities and distribution costs. In an uncapacitated facility location problem, facilities are opened at candidate locations without any capacity restrictions. The costs consist of fixed set-up costs when opening a facility and unit costs charged for each allocated unit (Erlenkotter, 1978).

Let $\mathcal{I} = \{1, ..., m\}$ be the set of candidate facility locations and $\mathcal{J} = \{1, ..., n\}$ the set of customers. Then, the demand of customer $j \in \mathcal{J}$ is D_j. The decision to open a facility $i \in \mathcal{I}$ leads to fixed set-up costs C_i. For $i \in \mathcal{I}$ and $j \in \mathcal{J}$, T_{ij} represents unit costs for serving customer j from facility i. Binary variables y_i reflect whether a facility $i \in \mathcal{I}$ is opened and continuous variables x_{ij} denote the number of product units transported from facility $i \in \mathcal{I}$ to customer $j \in \mathcal{J}$. The standard formulation of an uncapacitated facility location problem is given as (Fernández and Landete, 2019):

$$\min \sum_{i \in \mathcal{I}} C_i y_i + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} T_{ij} x_{ij}, \tag{1.1}$$

subject to:

$$\sum_{i \in \mathcal{I}} x_{ij} = D_j \qquad j \in \mathcal{J}, \tag{1.2}$$

$$x_{ij} \leq D_j y_i \qquad i \in \mathcal{I}, j \in \mathcal{J}, \tag{1.3}$$

$$x_{ij} \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, \tag{1.4}$$

$$y_i \in \{0, 1\} \qquad i \in \mathcal{I}. \tag{1.5}$$

Constraints (1.2) ensure that demand is satisfied, while constraints (1.3) guarantee that customers are not served from non-open facilities. Constraints (1.4) and (1.5) are the non-negativity and binary requirements. Note that in the optimal solution, each customer is served only from one facility. If some customers are served from more than one facility in the optimal solution, the transportation costs must be equal.

In contrast, the capacitated facility location problem considers both location and capacity of facilities needed to satisfy customer demand (Baumol and Wolfe, 1958). Due to capacity limitations, one customer can be served from more than one facility in the optimal solution. In the formulation of the capacitated facility location problem, constraints (1.3) are replaced as follows (Fernández and Landete, 2019):

$$\sum_{j \in \mathcal{J}} x_{ij} \leq Q_i y_i \qquad i \in \mathcal{I}, \tag{1.6}$$

where Q_i represents the capacity of facility $i \in \mathcal{I}$.

When considering different set-up costs for different facility sizes, the decision is extended to whether we should prefer many local facilities being close to customers' locations over a few large-scale centralized facilities. Local facilities benefit from low distribution costs, while large-scale facilities can profit from economies of scale. A solution to an optimisation model provides valuable decision support since it evaluates whether transportation costs dominate the economies of scale or vice versa.

1.3.2 Multi-period facility location

The main drawback of static facility location problems is that they consider only constant parameters and do not allow for dynamic parameter changes over time. However, as demand quantities and costs change over time, a model formulation considering time-dependent parameters and decision variables becomes necessary. A simple multi-period capacitated facility location problem can be formulated as Nickel and Saldanha da Gama (2019):

$$\min \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} C_{it} y_{it} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} T_{ijt} x_{ijt}, \tag{1.7}$$

subject to:

$$\sum_{i \in \mathcal{I}} x_{ijt} = D_{jt} \qquad j \in \mathcal{J}, t \in \mathcal{T}, \tag{1.8}$$

$$\sum_{j \in \mathcal{J}} x_{ijt} \leq Q_{it} y_{it} \qquad i \in \mathcal{I}, t \in \mathcal{T}, \tag{1.9}$$

$$x_{ijt} \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}, \tag{1.10}$$

$$y_{it} \in \{0, 1\} \qquad i \in \mathcal{I}, t \in \mathcal{T}. \tag{1.11}$$

This model extends the static formulation of the capacitated facility location problem by considering a planning horizon consisting of a set of time periods $\mathcal{T} = \{1, ..., \overline{T}\}$. Note that this simple model can be decomposed into time periods. For various formulations of dynamic facility location problems, see Nickel and Saldanha da Gama (2019).

In multi-period facility location problems, it is essential to consider the long-term impact of facility location decisions. Decisions that seem to be effective in the short-term horizon may disable future capacity expansion or lead to lock-in effects in the long term. Hence, the problem becomes a sequential decision process as all future steps depend on decisions taken earlier (Wesolowsky, 1973).

In a dynamic environment, capacity expansion and reduction are important steps to consider. The capability to change produced quantities might be necessary to satisfy time-varying demand. The overall capacity can be adjusted by opening new facilities or closing existing ones (Shulman, 1991). Another option is a modification of existing facilities in terms of capacity expansion or reduction (Jena et al., 2017). In the long term, capacity expansion can lead to considerable savings due to economies of scale. Similarly, capacity reduction may decrease the maintenance costs when production is down-scaled. Therefore, it is necessary to assess the initial investment and capacity adjustment costs with respect to future values of time-varying parameters. The ability to adjust the installed capacity impacts the initial investment decisions to a high degree. Facilities that can be easily adjusted can respond to demand changes with low additional costs, making them more attractive for investment.

1.3.3 Facility location under uncertainty

In deterministic problems, all inputs are available and known with certainty. In real-world problems, exact information about future parameters, such as demand and costs, is rarely available before investment decisions have to be taken. Traditionally, investment decisions are first-stage decisions as they have to be taken before uncertainty is disclosed. Demand allocation is then the second-stage decision since it depends on the realization of uncertain parameters (Birge and Louveaux, 2011).

Considering a finite set of scenarios $\mathcal{S}$ representing the uncertainty, the deterministic facility location model can be extended by scenario-indexed parameters. Let d_j^s be scenario dependent parameters representing demand of customer $j \in \mathcal{J}$ in scenario $s \in \mathcal{S}$, and T_{ij}^s unit costs for serving customer $j \in \mathcal{J}$ from facility $i \in \mathcal{I}$ in scenario $s \in \mathcal{S}$. Variable x_{ij}^s represents the number of product units transported from facility $i \in \mathcal{I}$ to customer $j \in \mathcal{J}$ in scenario $s \in \mathcal{S}$. The probability of a realization of a scenario $s \in \mathcal{S}$ is denoted p^s. A single-period two-stage stochastic uncapacitated facility location problem is formulated as (Correia and Saldanha da Gama, 2019):

$$\min \sum_{i \in \mathcal{I}} C_i y_i + \sum_{s \in \mathcal{S}} p^s \left(\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} T_{ij}^s x_{ij}^s \right), \tag{1.12}$$

subject to:

$$\sum_{i \in \mathcal{I}} x_{ij}^s = D_j^s \qquad j \in \mathcal{J}, s \in \mathcal{S}, \tag{1.13}$$

$$x_{ij}^s \leq D_j^s y_i \qquad i \in \mathcal{I}, j \in \mathcal{J}, s \in \mathcal{S}, \tag{1.14}$$

$$x_{ij}^s \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, s \in \mathcal{S}, \tag{1.15}$$

$$y_i \in \{0, 1\} \qquad i \in \mathcal{I}. \quad (1.16)$$

Note that the uncapacitated two-stage stochastic facility location problem has the property of relatively complete recourse, i.e., for every first-stage solution $(y_i, i \in \mathcal{I})$, there exists at least one feasible second-stage solution $(x^s_{ij}, i \in \mathcal{I}, j \in \mathcal{J}, s \in \mathcal{S})$ for all scenarios.

In the capacitated two-stage stochastic facility location problem, constraint (1.14) is replaced by:

$$\sum_{j \in \mathcal{J}} x^s_{ij} \leq Q_i y_i \qquad i \in \mathcal{I}, j \in \mathcal{J}, s \in \mathcal{S}. \quad (1.17)$$

Then however, the relatively complete recourse is not ensured which adds additional difficulty to this problem. A traditional way to deal with this infeasibility is to introduce penalties for unsatisfied demand. (Correia and Melo, 2019).

Traditionally, two-stage stochastic facility location problems have been formulated as single-period problems. When considering a multi-period formulation, the importance of including capacity adjustments increases since capacity adjustments can substantially improve flexibility in responding to demand changes. There is a crucial difference in whether capacity adjustments are considered first-stage or second-stage decisions. It can be shown that the ability to adjust the capacity in the second stage can lead to considerable savings compared to the case where capacity adjustments are first-stage decisions together with investment (Correia and Melo, 2021).

1.4 Related literature

We organize the related literature into three main groups. In Section 1.4.1, we focus on work related to the modelling of deterministic and stochastic facility location and capacity expansion problems. Within this section, literature related to economies of scale is discussed as well. Since some supply chain design problems have similar characteristics to facility location problems, we review hydrogen supply chain problems in Section 1.4.2. Finally, we discuss solution methods related to facility location problems in Section 1.4.3

1.4.1 Facility location problems

In this section, we first discuss deterministic multi-period facility location and capacity expansion problems. Then, we review the modelling of economies of scale in facility location problems, and finally, we discuss stochastic multi-period facility location and capacity expansion problems.

A review on deterministic multi-period facility location and capacity expansion problems, as well as on supply chain design and related solution methods can be found in Melo et al. (2006), Melo et al. (2009), Arabani and Farahani (2012), Nickel and Saldanha da Gama (2019), and Alarcon-Gerbier and Buscher (2022).

There are, in general, two main approaches to model capacity adjustments when considering modular capacities. First, capacity expansion is modelled as a building of additional facilities at the same location, while capacity reduction leads to facility closing (see, e.g., Shulman, 1991; Dias et al., 2007). Second, capacity expansion and reduction are modelled as a modification of an existing facility (see, e.g., Jena et al., 2015, 2016, 2017). Jena et al. (2015, 2016, 2017) also allows for facility closing and reopening later during the planning horizon. Capacitated facility location problems with continuous capacities can be seen in Hinojosa et al. (2008); Behmardi and Lee (2008); Torres-Soto and Üster (2011).

Having economies of scale in the production process enables taking cost advantage of up-scaling the production since a higher amount of products sharing the fixed costs results in lower unit costs (Haldi and Whitcomb, 1967). Holmberg (1994) uses a piecewise linear staircase cost function to model different production costs at different capacity levels. Correia and Captivo (2003) present a modular facility location problem that allows for the modelling of economies of scale since specific operational costs are defined for each facility size. Similar approach can be seen in Jena et al. (2015, 2016, 2017) and Correia and Melo (2021). Van den Broek et al. (2006) and Schütz et al. (2008) present a modelling approach for the continuous non-linear, non-convex and non-concave objective function. Their cost function can be seen as a combination of the staircase cost approach (Holmberg, 1994) together with capacity-dependent non-linear costs (Correia and Captivo, 2003).

Štádlerová and Schütz (2021) and Štádlerová et al. (2022b) belong to the category of multi-period facility location and capacity expansion papers with modular capacities where capacity expansion leads to facility modification. In these papers, the number of capacity expansions is limited, and capacity expansion only in terms of increasing the capacity level is allowed. The modelling of economies of scale for modular capacities presented in Correia and Captivo (2003) is extended by using specific piecewise-linear short-term production cost functions for each capacity level. Further, minimum production requirements for each capacity level are considered.

In recent years, single-period two-stage facility location problems have been extensively studied. An early review on facility location and supply chain problems with uncertain parameters can be seen in Owen and Daskin (1998) and Snyder (2006). Recent summaries on facility location and supply chain problems under uncertainty can be found in Govindan et al. (2017) and Correia and Saldanha da Gama (2019). However, there are relatively few papers dealing with two-stage stochastic multi-period facility location and capacity expansion problems (see, e.g., Zhuge et al., 2016; Correia and Melo, 2021). Correia and Melo (2021) present two-stage stochastic multi-period facility location and capacity expansion problem where binary capacity adjustments are second-stage decisions.

Šťádlerová et al. (2022a) and Šťádlerová et al. (2023) present a two-stage multi-period facility location and capacity expansion model for hydrogen infrastructure planning in Norway. Decisions related to binary capacity expansion in terms of increasing the capacity level are taken in the second stage together with decisions about capacity utilization and demand allocation. Similar to Šťádlerová and Schütz (2021) and Šťádlerová et al. (2022b), piecewise-linear convex short-term cost function for each capacity level and minimum production requirements based on technological limitations for electrolysis are considered.

1.4.2 Supply chain design problems

Some supply chain design problems have similar characteristics as facility location problems since they also consist of binary decisions related to the location of production facilities. Similarly, some two-stage stochastic supply chain problems consist of binary decisions related to facility location in the first stage and demand allocation problems in the second stage (Lucas et al., 2001). We refer to Li et al. (2019) for an overview of deterministic as well as stochastic hydrogen supply chain problems.

Almansoori and Shah (2009) study a multi-period hydrogen supply chain designed for Great Britain. Capacity expansion is not allowed. However, the authors consider hydrogen storage facilities. The results show, similar to Šťádlerová and Schütz (2021), that small facilities are opened at the beginning of the planning horizon when demand is relatively low. In later periods, when demand increases, more facilities with higher capacities are built. Myklebust et al. (2010) provide a case study from Germany and assess the optimal technology choice based on demand and input costs. The authors evaluate electrolysis and steam methane reforming with carbon capture. The decisive factor for electrolysis is electricity price, while the methane and CO_2 disposal price and demand level are the critical factors for steam methane reforming with carbon capture. Han et al. (2012) present a different approach since their model is formulated as a profit maximization model where facility location and capacity are given.

Šťádlerová and Schütz (2021) extend the previous formulations by considering capacity expansion as well. Further, the technology choice between electrolysis and steam methane reforming is a model decision. The results show that a decentralized solution consisting of small electrolysis facilities is preferred due to high distribution costs and low demand in the first time periods.

The hydrogen supply chain problem with uncertain demand is presented by Kim et al. (2008). The problem is formulated as a mixed-integer two-stage single-period supply chain design problem. In the first stage, decisions related to facility location and storage location are taken. In the second stage, demand allocation decisions are taken. A multi-period extension of Kim et al. (2008) can be found in Almansoori and Shah (2012) and Nunes et al. (2015). Dayhim et al. (2014) present the problem of minimizing expected daily costs of hydrogen demand considering emission, risk and consumption costs.

Similar to Nunes et al. (2015), Štádlerová et al. (2022a) also present a two-stage stochastic multi-period problem consisting of locating hydrogen facilities in the first stage. However, Štádlerová et al. (2022a) consider in addition capacity expansion in the second stage as a reaction to growing demand.

1.4.3 Solution methods

In this section, we first present solution methods for deterministic facility location problems. Then, we review solution methods for single-period as well as multi-period two-stage stochastic facility location problems.

Facility location and capacity expansion problems are in general hard to solve. For large instances, they may become intractable for commercial software. By introducing uncertainty, the complexity increases and the application of efficient solution methods becomes necessary to find a good feasible solution.

Lagrangian relaxation performs well for facility location problems. Relaxing the demand constraint and separating the relaxed problem into subproblems, one for each facility location, has been a popular choice in literature. Shulman (1991) shows that the optimisation problem for a single facility can be formulated as a shortest path problem and solved using a dynamic programming algorithm. The subgradient method based on Polyak (1969) is used to solve the Lagrangian dual problem. Jena et al. (2016, 2017) present a similar solution method, and they also compare the subgradient and bundle methods to update the Lagrangian multipliers showing that the bundle method has better convergence. Hinojosa et al. (2008) use Lagrangian relaxation for a problem with inventory constraints. Demand and flow conservation constraints have to be relaxed to obtain subproblems separable in facility location.

Another popular exact method is Benders decomposition. Here, the problem is decomposed into a master problem containing all binary variables and a subproblem with only continuous variables. The master problem provides an opening and expansion schedule, while the optimal demand allocation is solved in the subproblem. Castro et al. (2017) use Benders decomposition to solve a multi-period facility location and capacity expansion model. Torres-Soto and Üster (2011) compare Benders decomposition and Lagrangian relaxation on a facility location problem with possible relocation. The authors show that the problem and data structure are decisive for the performance of tested methods.

In recent years, heuristics have become a popular choice to solve facility location problems. However, they do not provide information about solution quality without a valid bound. Li et al. (2009) combine Lagrangian relaxation with tabu search heuristics on top to improve the upper bound. Tabu search is also used in Melo et al. (2012) to solve a facility relocation problem. A performance comparison of several heuristics for facility location problems can be found in Arostegui Jr et al. (2006) and for a facility location and capacity expansion problem in Silva et al. (2021).

The paper Šťádlerová et al. (2022b) contributes to the group of exact solution methods based on Lagrangian relaxation to solve multi-period facility location and capacity expansion problems. Unlike Shulman (1991) and Jena et al. (2016, 2017), Šťádlerová et al. (2022b) deal with a specific piecewise-linear convex short-term cost function for each capacity level as well as with minimum production requirements and limits on the number of capacity expansions.

Introducing uncertainty increases the complexity of the problem. When using the sample average approximation (SAA), the problem is solved repeatedly with a smaller number of scenarios to improve the tractability of the problem (Kleywegt et al., 2001). Nunes et al. (2015) use SAA to solve the hydrogen facility location problem since the problem can be solved only for fifteen scenarios. Šťádlerová et al. (2022a) extend the hydrogen facility location problem by allowing capacity expansion in the second stage. Due to binary variables in the second stage, the problem can be solved only for a maximum of ten scenarios. SAA is often used in combination with other solution methods to increase the number of scenarios in the sample and to further improve the quality of the solution (see, e.g., Santoso et al. (2005); Schütz et al. (2009); Li and Zhang (2018)).

Benders decomposition combined with SAA is presented in Santoso et al. (2005) to solve a supply chain design problem having only continuous variables in the second stage. Sherali and Zhu (2006) and Angulo et al. (2016) study Benders decomposition for problems with integer second stage.

Solution methods based on Lagrangian relaxation perform well for single-period two-stage stochastic facility location problems (Schütz et al., 2008). The paper Šťádlerová et al. (2023) contributes to the methodology of solving two-stage stochastic multi-period facility location and capacity expansion problems using Lagrangian relaxation. To the best of our knowledge, our work is the first to present a solution method based on Lagrangian relaxation for the multi-period facility location problem with uncertain demand and binary capacity expansion in the second stage.

1.5 Papers

The papers have been submitted to different scientific journals, having different layouts and requiring different referencing styles. For collecting the papers in this thesis, I have standardized the layout and the formatting of references. The content of the papers has not been modified.

Paper 1: Designing the hydrogen supply chain for maritime transportation in Norway

In paper 1, we study the problem of locating hydrogen production facilities for the maritime and land-based transportation sector in Norway. We present a multi-period facility location and capacity expansion model with modular capacities and specific non-linear production costs for each capacity level. The objective is to minimize the sum of investment, expansion, production, and distribution costs while satisfying the demand in each period. We present a model where the opening of new facilities is allowed during the whole planning horizon and compare it with a model where the opening of new facilities is possible only in the first time period. In both models, capacity expansion of existing facilities is allowed once during the planning horizon. We further analyze results for two demand scenarios: demand only from the maritime sector and demand from the maritime and land-based transportation sectors. The results show that the solution opens the same number of facilities independent of which demand scenario is used. However, there is a difference in the capacity level of the opened facilities. Considering a higher demand level, the importance of capacity expansion increases. The results further indicate that the initial demand is too low to build a steam methane reforming facility. Instead, only electrolysis facilities are opened.

This paper contributes to the modelling of multi-period facility location and capacity expansion problems. It further extends hydrogen supply chain design models by considering capacity expansion as well. However, the number of capacity expansions is limited. We present a model formulation with long-term investment and expansion costs separated from short-term production costs. The short-term production cost function is specific for each capacity level reflecting minimum production requirements as well. We analyze two models that differ in investment decision flexibility and the impact of investment flexibility on the number of expansions. In our model, the technology choice between electrolysis and steam methane reforming with carbon capture is a binary model decision. We address the real-world case of establishing the hydrogen infrastructure for maritime and road transportation in Norway. We demonstrate, similar to Keipi et al. (2018), that electrolysis is a more suitable technology for small scale production.

My contribution has been to formulate and implement the model, collect input data and perform numerical experiments. I have written large parts of the paper.

Co-author: Peter Schütz

The paper is published as: Štádlerová, Š., Schütz, P. (2021). Designing the Hydrogen Supply Chain for Maritime transportation in Norway. In: Mes, M., Lalla-Ruiz, E., Voß, S. (eds) *Computational Logistics.* ICCL 2021. Lecture Notes in Computer Science, volume 13004, 36-50. Springer, Cham

Paper 2: Multi-period facility location and capacity expansion with modular capacities and convex short-term costs

Paper 2 can be seen as a natural extension of Paper 1. Since we have seen that for larger instances, the problem is hard to solve with commercial software, we propose a solution method based on Lagrangian relaxation to solve a multi-period facility location and capacity expansion problem for large instances within a reasonable time. Our problem is characterized by minimum production requirements and convex piecewise-linear production costs for each capacity. The lower bound is formulated as a shortest path problem and solved using dynamic programming. We develop a greedy heuristic based on the lower-bound solutions to obtain a feasible solution. The results indicate that our solution method outperforms Gurobi in terms of run time and the advantage of our algorithm is more pronounced for larger instances. The results further show that our algorithm can find good solutions even for instances where Gurobi does not find any feasible solution within 24 hours.

This paper contributes to the group of exact solution methods based on Lagrangian relaxation to solve multi-period facility location and capacity expansion problems. Unlike Shulman (1991) and Jena et al. (2016, 2017), we deal with a specific piecewise-linear convex short-term cost function for each capacity level as well as with minimum production requirements and limits on the number of capacity expansions. We provide a comparison of our solution method based on Lagrangian relaxation, a restricted MIP approach (using Gurobi as a solver) utilizing the dual information from solving the Lagrangian dual, and Gurobi. We analyze the performance for instances that differ in size, demand level, and shape of the cost function. We show that only our solution method based on Lagrangian relaxation is capable of finding good solutions within a reasonable run time for all instances.

My contribution has been to formulate and implement the model, develop the solution method, collect input data and perform numerical experiments. I have written large parts of the paper.

Co-authors: Peter Schütz, Asgeir Tomasgard

The paper is submitted to an international journal.

Paper 3: Locating hydrogen production in Norway under uncertainty

Paper 3 is a stochastic extension of Paper 1. In Paper 1, we have considered a deterministic demand level. However, future demand estimates are highly uncertain, and the demand level has a crucial impact on binary decisions. Therefore, we study a problem where uncertainty in demand is considered.

We formulate our problem as a two-stage stochastic multi-period facility location and capacity expansion problem. In the first stage, decisions related to the opening of new facilities must be taken. In the second stage, existing facilities can be expanded. However, capacity expansion is allowed only once and only in terms of increasing the capacity level. Since the problem can be solved only for ten demand scenarios, we use sample average approximation to obtain good first-stage decisions. The solution to the stochastic problem results in lower installed capacity in the opening decisions compared to the expected value problem. The results further show that first-stage decisions using the expected value problem are infeasible for scenarios with low demand since minimum production requirements are violated. We further provide a managerial insight into the expected hydrogen costs in relation to capacity utilization.

This paper contributes to the modelling of two-stage stochastic multi-period facility location and capacity expansion problems. Following Correia and Melo (2021), we consider capacity expansion to be a recourse action in the second stage which considerably reduces the costs. We provide a general model formulation capable of incorporating minimum production requirements for each modular capacity level. Further, we define a specific convex piecewise linear short-term production costs function for each capacity level allowing for the modelling of production costs depending on capacity utilization.

My contribution has been to reformulate the model and extend the analysis provided by the co-authors in the Master thesis. I have implemented the model and the solution method and carried out new numerical experiments. I have written large parts of the paper.

Co-authors: Trygve Magnus Aglen, Andreas Hofstad, Peter Schütz

The paper is published as: Štádlerová, Š., Aglen, T.M., Hofstad, A., Schütz, P. (2022). Locating Hydrogen Production in Norway Under Uncertainty. In: de Armas, J., Ramalhinho, H., Voß, S. (eds) *Computational Logistics.* ICCL 2022. Lecture Notes in Computer Science, volume 13557, 306-321. Springer, Cham.

Paper 4: Using Lagrangian relaxation to locate hydrogen production facilities under uncertain demand: A case study from Norway

Paper 4 can be seen as an extension of Paper 2 and Paper 3. Since the stochastic problem can be solved only for ten scenarios using commercial software and the computational time is about five days, an efficient solution method is required to solve instances with more scenarios. Therefore, we extend the solution method from Paper 2 and present a solution method based on Lagrangian relaxation to solve the two-stage stochastic multi-period facility location and capacity expansion problem with uncertain demand. We formulate the relaxed problem as an expected shortest path problem and solve it using a dynamic programming algorithm. We develop a greedy heuristic based on the solution to the relaxed problem to solve the original problem. The results show that our solution method is capable of providing good results within a reasonable run time even for large real-world instances. This paper further provides a managerial insight into the solution structure considering different probabilistic demand distributions.

This paper contributes to the exact methodology of solving two-stage stochastic multi-period facility location and capacity expansion problems with binary expansion decisions in the second stage. These problems are new in the literature, and they have been solved only with a few scenarios (see, e.g., Correia and Melo, 2021; Štádlerová et al., 2022a). Our approach provides good results for real-world problems with up to 100 scenarios. To the best of our knowledge, our work is the first to present a solution method based on Lagrangian relaxation for the multi-period facility location problem with uncertain demand and binary capacity expansion in the second stage.

My contribution has been to formulate and implement the model, develop the solution method, collect input data and perform numerical experiments. I have written large parts of the paper.

Co-authors: Sanjay Dominik Jena, Peter Schütz

The paper is published as: Štádlerová, Š., Jena, S.D. & Schütz, P. Using Lagrangian relaxation to locate hydrogen production facilities under uncertain demand: a case study from Norway. *Computational Management Science* **20**, 10 (2023)

Bibliography

Aarskog, F. G., Danebergs, J., Strømgren, T., and Ulleberg, Ø. (2020). Energy and cost analysis of a hydrogen driven high speed passenger ferry. *International Shipbuilding Progress*, 67(1):97–123.

Alarcon-Gerbier, E. and Buscher, U. (2022). Modular and mobile facility location problems: A systematic review. *Computers & Industrial Engineering*, page 108734.

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

Almansoori, A. and Shah, N. (2012). Design and operation of a stochastic hydrogen supply chain network under demand uncertainty. *International Journal of Hydrogen Energy*, 37(5):3965–3977.

Andrenacci, S., Yejung, C., Raka, Y., Talic, B., and Colmenares-Rausseo, L. (2022). *Electrolysers towards EU MAWP 2023 targets and beyond*. Zenodo.

Angulo, G., Ahmed, S., and Dey, S. S. (2016). Improving the integer l-shaped method. *INFORMS Journal on Computing*, 28(3):483–499.

Arabani, A. B. and Farahani, R. Z. (2012). Facility location dynamics: An overview of classifications and applications. *Computers & Industrial Engineering*, 62(1):408–420.

Arostegui Jr, M. A., Kadipasaoglu, S. N., and Khumawala, B. M. (2006). An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems. *International Journal of Production Economics*, 103(2):742–754.

Baumol, W. J. and Wolfe, P. (1958). A warehouse-location problem. *Operations research*, 6(2):252–263.

Behmardi, B. and Lee, S. (2008). Dynamic multi-commodity capacitated facility location problem in supply chain. In Fowler, J. and Mason, S., editors, *Proceedings of the 2008 Industrial Engineering Research Conference*, pages 1914–1919. Institute of Industrial and Systems Engineers (IISE).

Birge, J. R. and Louveaux, F. (2011). *Introduction to stochastic programming.* Springer Science & Business Media, New York, 2 edition.

Castro, J., Nasini, S., and Saldanha da Gama, F. (2017). A cutting-plane approach for large-scale capacitated multi-period facility location using a specialized interior-point method. *Mathematical Programming*, 163(1-2):411–444.

Correia, I. and Captivo, M. E. (2003). A Lagrangean heuristic for a modular capacitated location problem. *Annals of Operations Research*, 122(1):141–161.

Correia, I. and Melo, T. (2019). Dynamic facility location problem with modular capacity adjustments under uncertainty. Technical report, Schriftenreihe Logistik der Fakultät für Wirtschaftswissenschaften der htw saar.

Correia, I. and Melo, T. (2021). Integrated facility location and capacity planning under uncertainty. *Computational and Applied Mathematics*, 40(5):1–36.

Correia, I. and Saldanha da Gama, F. (2019). Facility location under uncertainty. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, pages 185–213. Springer.

Danebergs, J. and Aarskog, F. G. (2020). Future compressed hydrogen infrastructure for the domestic maritime sector. IFE/E-2020/006, Halden, Norway.

Dayhim, M., Jafari, M. A., and Mazurek, M. (2014). Planning sustainable hydrogen supply chain infrastructure with uncertain demand. *International Journal of Hydrogen Energy*, 39(13):6789–6801.

Dias, J., Captivo, M. E., and Clímaco, J. (2007). Dynamic location problems with discrete expansion and reduction sizes of available capacities. *Investigação Operacional*, 27(2):107–130.

DNV GL (2019). Produksjon og bruk av hydrogen i Norge. Rapport 2019-0039, Oslo, Norway, (in Norwegian).

DNV GL (2020). Energy transition outlook 2020: A global and regional forecast to 2050. Oslo, Norway.

Erlenkotter, D. (1978). A dual-based procedure for uncapacitated facility location. *Operations Research*, 26(6):992–1009.

Fernández, E. and Landete, M. (2019). Fixed-charge facility location problems. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, pages 67–98. Springer, Cham.

Fridstrøm, L., Tomasgard, A., Eskeland, G. S., Espegren, K. A., Rosenberg, E., Helgesen, P. I., Lind, A., Ryghaug, M., Berg, H. B., Walnum, H. J., Ellingsen, L., and Graabak, I. (2018). Decarbonization of transport, a position paper prepared by FME MoZEES and FME CenSES. ISBN 978-82-93198-25-3.

Govindan, K., Fattahi, M., and Keyvanshokooh, E. (2017). Supply chain network design under uncertainty: A comprehensive review and future research directions. *European Journal of Operational Research*, 263(1):108–141.

Haldi, J. and Whitcomb, D. (1967). Economies of scale in industrial plants. *Journal of Political Economy*, 75(4, Part 1):373–385.

Han, J.-H., Ryu, J.-H., and Lee, I.-B. (2012). Modeling the operation of hydrogen supply networks considering facility location. *International Journal of Hydrogen Energy*, 37(6):5328–5346.

Hinojosa, Y., Kalcsics, J., Nickel, S., Puerto, J., and Velten, S. (2008). Dynamic supply chain design with inventory. *Computers & Operations Research*, 35(2):373–391.

Hirth, M., Hove, K., Janzen, D., Eide, P., Helland, P., Østvik, I., Ryberg, T., and Ødegard, J. (2019). Norwegian future value chains for liquid hydrogen. NCE Maritime CleanTech, Report liquid hydrogen 2019, Stord, Norway.

Holmberg, K. (1994). Solving the staircase cost facility location problem with decomposition and piecewise linearization. *European Journal of Operational Research*, 75(1):41–61.

IEA (2022a). Global hydrogen review 2022. *International Energy Agency*, pages 1–284.

IEA (2022b). Hydrogen: Energy system overview. *International Energy Agency*. Tracking report, Paris.

IRENA (2020). Green hydrogen: A guide to policy making. *International Renewable Energy Agency*.

Jakobsen, D. and Åtland, V. (2016). Concepts for large scale hydrogen production. Master's thesis, Department of Energy and Process Engineering, NTNU, Trondheim, Norway.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2015). Dynamic facility location with generalized modular capacities. *Transportation Science*, 49(3):484–499.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2016). Solving a dynamic facility location problem with partial closing and reopening. *Computers & Operations Research*, 67:143–154.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2017). Lagrangian heuristics for large-scale dynamic facility location with generalized modular capacities. *INFORMS Journal on Computing*, 29(3):388–404.

Kall, P. and Wallace, S. W. (1994). *Stochastic programming*, volume 6. Springer.

Keipi, T., Tolvanen, H., and Konttinen, J. (2018). Economic analysis of hydrogen production by methane thermal decomposition: Comparison to competing technologies. *Energy Conversion and Management*, 159:264–273.

Kim, J., Lee, Y., and Moon, I. (2008). Optimization of a hydrogen supply chain under demand uncertainty. *International Journal of Hydrogen Energy*, 33(18):4715–4729.

Kleywegt, A. J., Shapiro, A., and Homem-de Mello, T. (2001). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

Li, J., Chu, F., and Prins, C. (2009). Lower and upper bounds for a capacitated plant location problem with multicommodity flow. *Computers & Operations Research*, 36(11):3019–3030.

Li, L., Manier, H., and Manier, M.-A. (2019). Hydrogen supply chain network design: An optimization-oriented review. *Renewable and Sustainable Energy Reviews*, 103:342–360.

Li, X. and Zhang, K. (2018). A sample average approximation approach for supply chain network design with facility disruptions. *Computers & Industrial Engineering*, 126:243–251.

Lucas, C., MirHassani, S., Mitra, G., and Poojari, C. (2001). An application of lagrangian relaxation to a capacity planning problem under uncertainty. *Journal of the Operational Research Society*, 52(11):1256–1266.

Mäkitie, T., Hanson, J., Steen, M., Hansen, T., and Andersen, A. D. (2020). The sectoral interdependencies of low-carbon innovations in sustainability transitions. FME NTRANS Working paper 01/20, Trondheim, Norway.

Melo, M., Nickel, S., and Saldanha da Gama, F. (2012). A tabu search heuristic for redesigning a multi-echelon supply chain network over a planning horizon. *International Journal of Production Economics*, 136(1):218–230.

Melo, M. T., Nickel, S., and Saldanha da Gama, F. (2006). Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research*, 33(1):181–208.

Melo, M. T., Nickel, S., and Saldanha da Gama, F. (2009). Facility location and supply chain management–a review. *European Journal of Operational Research*, 196(2):401–412.

Myklebust, J., Holth, C., Tøftum, L. E. S., and Tomasgard, A. (2010). Optimizing investments for hydrogen infrastructure in the transport sector. In *Techno-economic modelling of value chains based on natural gas:with consideration of CO_2 emissions*, pages 27–70. Doctoral thesis at NTNU 2010:83, Department of Industrial Economics and Technology Management, Trondheim, Norway.

NEL Hydrogen (2015). Efficient electrolysers for hydrogen production. `http://wpstatic.idium.no/www.nel-hydrogen.com/2015/03/Efficient_Electrolysers_for_Hydrogen_Production.pdf/`. last accessed 05.02.2021.

Nemhauser, G. L. and Wolsey, L. A. (1999). *Integer and combinatorial optimization*, volume 55. John Wiley & Sons, Hoboken.

Nickel, S. and Saldanha da Gama, F. (2019). Multi-period facility location. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, pages 303–326. Springer, Cham.

Nunes, P., Oliveira, F., Hamacher, S., and Almansoori, A. (2015). Design of a hydrogen supply chain with uncertainty. *International Journal of Hydrogen Energy*, 40(46):16408–16418.

Ocean Hyway Cluster (2020). 2030 hydrogen demand in the Norwegian domestic maritime sector. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Owen, S. H. and Daskin, M. S. (1998). Strategic facility location: A review. *European Journal of Operational Research*, 111(3):423–447.

Polyak, B. T. (1969). Minimization of unsmooth functionals. *USSR Computational Mathematics and Mathematical Physics*, 9(3):14–29.

Samferdselsdepartementet (2021). Nasjonal transportplan 2022-2033. `https://www.regjeringen.no/no/dokumenter/meld.-st.-20-20202021/id2839503/?ch=1`. last accessed 09.03.2022,(in Norwegian).

Santoso, T., Ahmed, S., Goetschalckx, M., and Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115.

Schütz, P., Stougie, L., and Tomasgard, A. (2008). Stochastic facility location with general long-run costs and convex short-run costs. *Computers & Operations Research*, 35(9):2988–3000.

Schütz, P., Tomasgard, A., and Ahmed, S. (2009). Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199(2):409–419.

Sherali, H. D. and Zhu, X. (2006). On solving discrete two-stage stochastic programs having mixed-integer first-and second-stage variables. *Mathematical Programming*, 108(2):597–616.

Shulman, A. (1991). An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations Research*, 39(3):423–436.

Silva, A., Aloise, D., Coelho, L. C., and Rocha, C. (2021). Heuristics for the dynamic facility location problem with modular capacities. *European Journal of Operational Research*, 290(2):435–452.

Snyder, L. V. (2006). Facility location under uncertainty: a review. *IIE Transactions*, 38(7):547–564.

Štádlerová, Š., Aglen, T. M., Hofstad, A., and Schütz, P. (2022a). Locating hydrogen production in norway under uncertainty. In Ramalhinho, H., De Armas, J., and Voß, S., editors, *Computational Logistics*, volume 13557, pages 306–321. Springer, Cham.

Štádlerová, Š., Jena, S. D., and Schütz, P. (2023). Using Lagrangian relaxation to locate hydrogen production facilities under uncertain demand: a case study from Norway. *Computational Management Science*, 20(10).

Štádlerová, Š. and Schütz, P. (2021). Designing the hydrogen supply chain for maritime transportation in Norway. In Mes, M., Lalla-Ruiz, E., and Voß, S., editors, *Computational Logistics*, volume 13004, pages 36–50. Springer, Cham.

Štádlerová, Š., Schütz, P., and Tomasgard, A. (2022b). Multi-period facility location and capacity expansion with modular capacities and convex short-term costs. Working paper, Department of Industrial Economics and Technology Management, NTNU, Norway.

Torres-Soto, J. E. and Üster, H. (2011). Dynamic-demand capacitated facility location problems with and without relocation. *International Journal of Production Research*, 49(13):3979–4005.

Van den Broek, J., Schütz, P., Stougie, L., and Tomasgard, A. (2006). Location of slaughterhouses under economies of scale. *European Journal of Operational Research*, 175(2):740–750.

Wesolowsky, G. O. (1973). Dynamic facility location. *Management Science*, 19(11):1241–1248.

Zhuge, D., Yu, S., Zhen, L., and Wang, W. (2016). Multi-period distribution center location and scale decision in supply chain network. *Computers & Industrial Engineering*, 101:216–226.

Paper I

Designing the hydrogen supply chain for maritime transportation in Norway

Šárka Štádlerová, Peter Schütz

Published in Lecture Notes in Computer Science

Paper 1
Designing the hydrogen supply chain for maritime transportation in Norway

Abstract

We study the problem of locating hydrogen facilities for the maritime transportation sector in Norway. We present a multi-period model with capacity expansion to obtain optimal investment and expansion decisions and to choose optimal production quantities and distribution solutions. The objective is to minimize the sum of investment, expansion, production, and distribution costs while satisfying the demand in each period. Hydrogen production costs are subject to economies of scale which causes non-linearity in the objective function. We model long-term investment and expansion costs separately from short-term production costs. The short-term production costs depend on the installed capacity and production quantities. We analyze two models that differ in investment decision flexibility and two demand scenarios: demand only from the maritime sector and demand from the whole transportation sector in Norway. The results show that the scenario with higher demand does not lead to a higher number of built facilities due to the economies of scale. The model with higher flexibility leads to higher capacity utilization in the first periods and thus significantly lower production costs. The results further indicate that the initial demand is too low to build a steam methane reforming facility, instead only electrolysis facilities are built in both scenarios and both models.

Keywords: Facility location, Capacity expansion, Hydrogen supply chain

2.1 Introduction

Emission reduction in the transportation sector is a crucial step in order to meet the emission targets set in the Paris agreement on climate change. In 2015, the Norwegian parliament decided that CO_2 emissions must be decreased by at least 40% (compared to 1990) towards 2030 in an attempt to reach the targets of the Paris agreement. As a consequence of this ambitious decision, fossil fuels have to be replaced by alternative zero-emission fuels. The use of hydrogen fuel cells is considered as one way to decarbonize the transport sector and to decrease the emission of greenhouse gases (GHG) (Fridstrøm et al., 2018).

In 2017, the transport sector in Norway was responsible for emitting 15.8 mill. tons CO_2, accounting for 23% of all CO_2 emissions (Aarskog et al., 2020). CO_2 emissions from domestic inland water and coastal transport in Norway accounted for 8.7% of emissions from the transport sector in 2018. Introducing zero-emission fuels such as hydrogen in maritime transportation can therefore considerably reduce emissions of CO_2. However, limited experience with hydrogen as fuel and uncertainty about hydrogen availability may affect the smoothness of the transition to hydrogen fuels (Mäkitie et al., 2020). One way to create an initial demand for hydrogen is to require that high-speed passenger ferries and car ferries have to use hydrogen as fuel when public transport contracts are renewed. In general, demand for hydrogen is expected to increase in the years to come and the production infrastructure has to adjust to this growth (DNV GL, 2019). As such, the infrastructure needed to cover demand from the maritime sector can help ensuring a stable hydrogen supply also for other transportation sectors in Norway (Fridstrøm et al., 2018).

The two most relevant hydrogen production technologies for Norway are electrolysis (EL) and steam methane reforming with carbon capture (SMR+) (Hirth et al., 2019). While electrolysis is a more profitable technology in small-scale production (50–5,000Nm^3/h), SMR+ is more favourable when producing large quantities of hydrogen (50,000–100,000Nm^3/h). Scaling up the production results in lower average costs, leading to economies of scale. This property is significant for SMR+, but it also applies to electrolysis (Keipi et al., 2018). Figure 2.1 shows the economies of scale in the long-term hydrogen cost function. Note that the cost-axis uses a logarithmic scale.

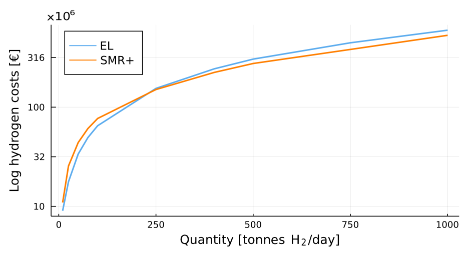


Figure 2.1: Long-term hydrogen costs

In this paper, we study the problem of how to design the hydrogen supply chain for maritime transportation in Norway. The problem consists of investment and expansion decisions, production quantities, and distribution solutions. It belongs to the category of facility location problems with capacity expansion. An early review of pioneering papers dealing with capacity expansion can be found in Luss (1982). Shulman (1991) and Dias et al. (2007) study a multi-period plant location problem with discrete expansion where a plant is modelled as a set of facilities in the same location. Capacity expansion is achieved by building an additional facility and the facility size must be chosen from a finite set of capacities. The production costs are defined for each facility and depend only on facility type and quantity produced in the facility. Behmardi and Lee (2008) study a multi-period multi-commodity capacitated facility location problem with capacity expansion and relocation. The modelling approach differs from previous papers as Behmardi and Lee (2008) work with dummy locations to relocate capacity. The dummy locations are used for modelling purposes to shift the capacity. Customers can only be served from real facilities. Torres-Soto and Üster (2011) present a comparison of multi-period facility location problems with growing demand where opening and closing decisions are allowed at any time during the planning horizon. Jena et al. (2015) introduce a multi-period facility location model with a capacity expansion, reduction, and the option to temporarily close the facility. In their work, capacity expansion is modelled by the modification of existing facilities. Jena et al. (2016) present a facility location problem with modular capacities where capacity expansion, as well as partial closing and reopening, are allowed. An extension of their model is published in Jena et al. (2017) where also facility relocation is allowed. Castro et al. (2017) present a large-scale capacitated multi-period facility location model where a set of capacitated facilities is progressively built during the planning horizon and simultaneously a maximum amount of operating facilities in each period is specified.

Facility location and supply chain design problems with a focus on hydrogen infrastructure are discussed in Almansoori and Shah (2009), Myklebust et al. (2010) and Han et al. (2012). In the work by Almansoori and Shah (2009), a multi-period hydrogen supply chain for Great Britain is studied. However, in their work, expansion is not allowed. Myklebust et al. (2010) present a case study from Germany and study the impact of demand and input costs on the optimal technology choice. Han et al. (2012) present a different approach where an optimization model for the hydrogen supply chain with given production capacities is considered.

Economies of scale cause non-linear production costs. Several approaches for how to incorporate non-linear production costs in facility location problems have been published in the literature. Holmberg (1994) introduces a piecewise linear staircase cost function that enables to model different production costs at different capacity levels. Correia and Captivo (2003) present the modular capacitated facility location model and emphasize the advantage of the modular formulation as it enables to take economies of scale into consideration. They separate investment and operational costs and provide different unit operational costs for each facility size. Van den Broek et al. (2006) study facility location problem with non-linear, non-convex, and non-concave objective function. They follow the idea of non-linear costs depending on installed capacity as presented in Correia and Captivo (2003) however, they introduce a linear staircase cost approximation. The approach presented in Van den Broek et al. (2006) can capture economies as well as diseconomies of scale.

For more examples of facility location and supply chain design see the excellent reviews by Melo et al. (2009), and Arabani and Farahani (2012). Review on multi-period facility location problems can be found in Nickel and Saldanha da Gama (2019).

In this paper, we investigate the impact of demand and decision flexibility on the optimal design of the hydrogen infrastructure for maritime transportation in Norway. In particular, we study where to locate hydrogen production facilities, which capacity and production technology to install, and which period to choose for investment and expansion.

We distinguish between long-term costs and short-term costs. Long-term costs consist of investment and expansion costs, while the short-term costs are given as production costs, representing capital expenses (CAPEX) and operational expenses (OPEX) respectively.

The investment and expansion represent the long-term decision because a built facility cannot be closed down during the planning horizon. The short-term production costs depend on installed capacity and its utilization. We allow the production rate to deviate from the installed capacity, allowing for a more flexible production schedule. However, deviating from the installed capacity leads to increasing unit costs (Schütz, 2009). We carry out our analysis using two models and two demand scenarios. In the first model, opening new facilities is allowed during the whole planning horizon, while in the second model, opening facilities is restricted to the first period. In the first demand scenario, we assume demand only from the maritime sector, while in the second scenario, demand from the whole transportation sector in Norway is considered.

The remainder of this paper is organized as follows: in Section 2.2, we provide a mathematical formulation of the dynamic facility location problem with capacity expansion. Case description and computational results are discussed in Sections 2.3 and 2.4, respectively. Conclusion is presented in Section 2.5.

2.2 The mathematical programming model

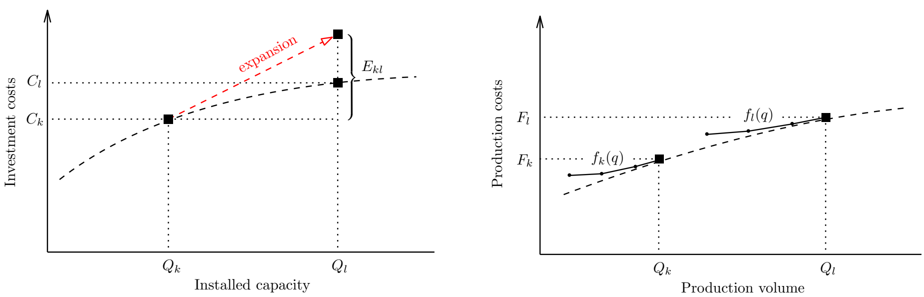
We formulate our problem as a multi-period facility location problem with capacity expansion. The goal is to determine the optimal strategy for opening and expanding hydrogen production facilities such that demand is satisfied. Closing facilities is not allowed. The objective is to minimize the discounted sum of investment and expansion costs, production costs, and distribution costs.

We provide two models for our multi-period facility location problem with non-linear objective function and capacity expansion. In the first model, investing in a new facility is allowed in each period, while in the second model, the initial investment can only be made in the first period. In both models, capacity expansion is allowed for each facility once during the planning horizon, and technology change is not permitted. We assume that the cost functions are independent of selected locations and investment time. Each technology is characterized by its own cost function. However, the general properties described in Subsection 2.2.1 apply to both considered technologies. The mathematical formulation is then presented in Subsection 2.2.2.

2.2.1 Modelling approach

We model investment decisions as a choice from a discrete set of available capacities similar to Correia and Captivo (2003). Capacity expansion here means modifying an existing facility and is modelled as a discrete jump between available capacities. This approach is also used in Jena et al. (2016).

To model the cost of investing, expanding, and operating facilities, we separate the long-term investment and expansion costs from the short-term production costs. Each installed capacity has its own short-term production cost function. We model the short-term production costs as a piecewise linear, convex function. This is similar to the approach presented in Schütz et al. (2008). From the point of view of short-term production costs, higher utilization of smaller capacity is always more favourable than smaller utilization of higher installed capacity.



(a) Investment and expansion costs (b) Production costs

Figure 2.2: Short-term and long-term costs

Expanding capacity implies an additional investment as well as switching over to a new short-term production cost function. Figure 2.2a illustrates our approach for modelling the expansion of facilities. Let Q_k be the initially installed capacity and C_k the corresponding investment costs. The expansion costs of expanding from capacity Q_k to capacity Q_l are denoted as E_{kl}. As $C_k + E_{kl} > C_l$. Investing in a smaller facility and expanding to a larger capacity is more expensive than opening the bigger facility right away.

Due to separating the long-term investment and expansion costs from the short-term production costs, expansion implies moving from one short-term production cost function to another. An example of this can be seen in Figure 2.2b. Before expanding the facility from capacity Q_k to capacity Q_l, the production cost function $f_k(q)$ applies, whereas function $f_l(q)$ is valid after the expansion has taken place.

2.2.2 Mathematical formulation

Let us first introduce the following notation:

Sets

$\mathcal{B}$ Set of breakpoints of the short-term cost function
$\mathcal{I}$ Set of possible facility locations
$\mathcal{J}$ Set of customer ports
$\mathcal{K}$ Set of available discrete capacities
$\mathcal{P}$ Set of periods
$\mathcal{T}$ Set of available production technologies

Parameters and coefficients

C_{ikt}	investment costs in location i, for point k of capacity function, and technology t;
D_{jp}	demand in port j in period p;
E_{klt}	costs of expansion from capacity in point k to capacity in point l for technology t;
F_{bkt}	costs at breakpoint b of the short-term cost function given for capacity k and for technology t;
L_{ijp}	1 if demand at location j can be served from facility i in period p, 0 otherwise;
Q_{bkt}	production volume at breakpoint b of the short-term cost function, for capacity point k and technology t;
T_{ijp}	transportation costs from facility i to customer j in period p;
y_{iklt0}	initial facility variable;
δ_p	discount factor in period p;
τ_p	length of time period p in years;

Decision variables

x_{ijp} amount of customer demand at location j satisfied from facility i in period p;

y_{ikltp} 1 if facility is opened in location i in period p, with originally installed capacity k, operated capacity l, and technology t, 0 otherwise;

μ_{biltp} weight of breakpoint b at location i for capacity point k and technology t in period p.

We present a multi-period model where investment and expansion decisions are allowed during the whole planning horizon. The changes in formulation needed for the first-period model are presented at the end of this section. The problem is given as:

$$\begin{aligned}
\min \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}}\sum_{l\in\{l\geq k:l\in\mathcal{K}\}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} \delta_p C_{ikt}\left(y_{ikltp}-y_{iklt(p-1)}\right)+ \\
\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}}\sum_{l\in\{l>k:l\in\mathcal{K}\}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} \delta_p E_{klt}(y_{ikltp}-y_{iklt(p-1)})+ \\
\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{p\in\mathcal{P}} \delta_p\tau_p T_{ijp}x_{ijp}+ \\
\sum_{b\in\mathcal{B}}\sum_{i\in\mathcal{I}}\sum_{l\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}} \delta_p\tau_p F_{blt}\mu_{biltp},
\end{aligned} \tag{2.1}$$

subject to:

$$\sum_{k\in\mathcal{K}}\sum_{l\in\{l\geq k:l\in\mathcal{K}\}}\sum_{t\in\mathcal{T}} y_{ikltp}\leq 1, \qquad p\in\mathcal{P}, \tag{2.2}$$

$$\sum_{l\in\{l\geq k:l\in\mathcal{K}\}} y_{ikltp}\geq \sum_{l\in\{l\geq k:l\in\mathcal{K}\}} y_{iklt(p-1)}, \quad i\in\mathcal{I},k\in\mathcal{K},t\in\mathcal{T},p\in\mathcal{P}, \tag{2.3}$$

$$y_{ikltp}-y_{iklt(p-1)}\geq 0, i\in\mathcal{I},k\in\mathcal{K},l\in\{l>k:l\in\mathcal{K}\},t\in\mathcal{T},p\in\mathcal{P}, \tag{2.4}$$

$$\sum_{b\in\mathcal{B}}\mu_{biltp}=\sum_{k\in\mathcal{K}} y_{ikltp}, \qquad i\in\mathcal{I},l\in\mathcal{K},t\in\mathcal{T},p\in\mathcal{P}, \tag{2.5}$$

$$\sum_{j\in\mathcal{J}} x_{ijp}=\sum_{b\in\mathcal{B}}\sum_{l\in\mathcal{K}}\sum_{t\in\mathcal{T}} Q_{blt}\mu_{biltp}, \qquad i\in\mathcal{I},p\in\mathcal{P}, \tag{2.6}$$

$$\sum_{i \in \mathcal{I}} x_{ijp} = D_{jp}, \qquad j \in \mathcal{J}, p \in \mathcal{P}, \quad (2.7)$$

$$x_{ijp} \leq L_{ijp} D_{ip}, \qquad i \in \mathcal{I}, j \in \mathcal{J}, p \in \mathcal{P}, \quad (2.8)$$

$$y_{ikltp} \in \{0, 1\}, \qquad i \in \mathcal{I}, k \in \mathcal{K}, l \in \{l \geq k : l \in \mathcal{K}\} t \in \mathcal{T}, p \in \mathcal{K}, \quad (2.9)$$

$$x_{ijp} \geq 0, \qquad i \in \mathcal{I}, j \in \mathcal{J}, p \in \mathcal{P}, \quad (2.10)$$

$$\mu_{biltp} \geq 0, \qquad b \in \mathcal{B}, i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, p \in \mathcal{P}. \quad (2.11)$$

The objective function (2.1) is the discounted sum of investment costs, expansion costs, distribution costs, and production costs. Restrictions (2.2) guarantee that only one facility can be opened at the given location. Constraints (2.3) ensure that a facility can expand but cannot be closed. Capacity expansion is allowed only once during the planning horizon. The variable y_{ikltp} contains information about the initially installed capacity k as well as the capacity l at which it is currently operated. After expansion, the operated capacity l is higher than the installed capacity k. Inequalities (2.4) ensure that capacity index l can change only once. Equations (2.5) ensure that production is allocated only to opened facilities and that the short-term production cost function depends on operated capacity. Equations (2.6) express the requirement that the whole production has to be distributed to customers. Equations (2.7) ensure demand satisfaction, while constraints (2.8) specify if customer j can be served from facility i. Restrictions (2.9) - (2.11) are the binary and non-negativity requirements.

In our second model, a facility can only be opened in the first period. Expansion is still allowed in later periods. In this model, constraint (2.12) replaces constraint (2.3):

$$\sum_{l \in \{l \geq k : l \in \mathcal{K}\}} y_{ikltp} = y_{ikkt1}, \qquad i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, p \in \mathcal{P}. \quad (2.12)$$

The rest of the model is identical to the first model.

2.3 Case study

In this section, we present the input data for the problem of designing the Norwegian hydrogen supply chain for maritime transportation. We include 17 candidate locations for hydrogen facilities on the Norwegian west coast. The candidate locations for hydrogen production are obtained from the interactive map set up by Ocean Hyway Cluster (2020b).

We consider two hydrogen production technologies: EL and SMR+. We approximate the facility capacity by 8 discrete points for EL and 7 points for SMR+. The discrete points are given in Table 2.1. We use the same discretization of capacity for both technologies, but we do not consider SMR+ for the smallest capacity. In Table 2.1, we provide facility investment costs and production costs per kilogram at the discrete capacity points. Note that with decreasing utilization, the production costs per unit increase (Ulleberg and Hancke, 2020).

Discrete capacity	1	2	3	4	5	6	7	8
Capacity [tonnes/day]	0.6	3.1	6.2	12.2	30.3	61.0	151.5	304.9
Investment EL [mill. €]	1.4	6.0	11.2	20.5	46.5	87.2	197.7	371.5
Investment SMR^+ [mill. €]	-	23.9	39.9	65.2	127.7	204.3	402.1	709.2
Production EL [€/kg]	1.95	1.61	1.53	1.45	1.43	1.42	1.40	1.38
Production SMR^+ [€/kg]	-	1.91	1.61	1.42	1.28	1.18	1.04	1.00

Table 2.1: Investment and production costs for EL and SMR+ at discrete capacity points

The production rate for an EL facility can vary between $20 - 100\%$ of the installed capacity (NEL Hydrogen, 2015). We define a piecewise linear, convex short-term production costs for each discrete capacity. We approximate the short-term production costs by a piecewise linear function with breakpoints at $20\%, 50\%, 80\%$ and 100% of installed production quantity. For simplification, we use the same production rates for SMR+. We use the model by Jakobsen and Åtland (2016) for calculating investment and short-term production costs for electrolysis and SMR+.

We calculate the expansion costs as the difference between the investment costs of opening two facilities with different capacities plus an additional mark-up. We assume the mark-up for expansion to be 10% of the difference in investment costs.

We derive the costs of distributing one kilogram of hydrogen for one kilometer for distances up to 800 km from Danebergs and Aarskog (2020). To obtain the costs for distributing up to 1000 km, we extrapolate the distribution cost function. The distribution costs per kilometer and kilogram hydrogen are then valid for the appropriate interval as shown in Table 2.2. If a customer is located in the same municipality as a facility, we assume zero distribution costs. We set the distance limit between production facility and customer to 1000 km. Hydrogen distribution over 1000 km is suitable for pipelines. However, pipelines are not considered relevant for Norway (Damman et al., 2020).

Distance [km]	1-50	51-100	101-200	201-400	401-800	801-1000
Costs	0.00498	0.00426	0.00390	0.00372	0.00363	0.00360

Table 2.2: Hydrogen distribution costs in [€/km/kg H_2]

We use two demand scenarios where hydrogen demand is increasing during the planning horizon (see Figure 2.3). In the maritime sector, demand moderately increases until period 11. In period 11, the coastal route Bergen-Kirkenes starts to operate on hydrogen fuels which causes a significant increase in demand. Until period 3, there is no difference between the two demand scenarios. In the whole transportation sector, the main demand growth is in periods 4 and 9 which corresponds to years 2025 and 2030. These dates represent two strategic phases for hydrogen transition in heavy transport and long-distance bus transport (DNV GL, 2019).

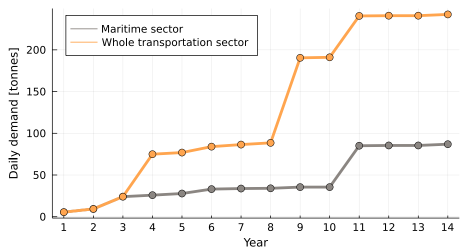


Figure 2.3: Development of hydrogen demand during the planning horizon

- Maritime: high-speed passenger ferries, car ferries, and coastal route Bergen-Kirkenes, (Aarskog and Danebergs, 2020), (Ocean Hyway Cluster, 2020a)
- All transportation: maritime sector plus road traffic and railway sector, (DNV GL, 2019)

Aarskog and Danebergs (2020) and Ocean Hyway Cluster (2020a) present high-speed passenger ferry and car ferry routes that are relevant for hydrogen fuel as well as their bunkering locations. They list 51 relevant customer locations for the maritime sector and assume that new contracts for public transportation services will require a zero-emission solution and that hydrogen will be selected as fuel. For the whole transportation sector, the list of customers is extended to 70 locations and consists of bunkering ports and several inland locations relevant for hydrogen consumption in road traffic and the railway sector.

In our case, we assume the discounting interest rate to be zero. Thus, the discount factor δ_p is equal to one in each period.

2.4 Computational results

The model is implemented in Mosel and solved with Xpress Optmizer Version 36.01.10. All calculations were run on a laptop with a Intel(R) Core(TM) i7-10510U CPU @ 1.80GHz processor and 16GB RAM.

A summary of the main results of both demand scenarios and both models can be found in Table 2.3. We provide the main characteristics of the built infrastructure as the number of built facilities and the number of expansions. Total capacity and average size refer to the installed capacity and average facility size in the last period. The total costs represent the sum of investment, expansion, production, and distribution costs. The average hydrogen costs are calculated over the entire planning horizon average. Note that the chosen technology is electrolysis in all cases.

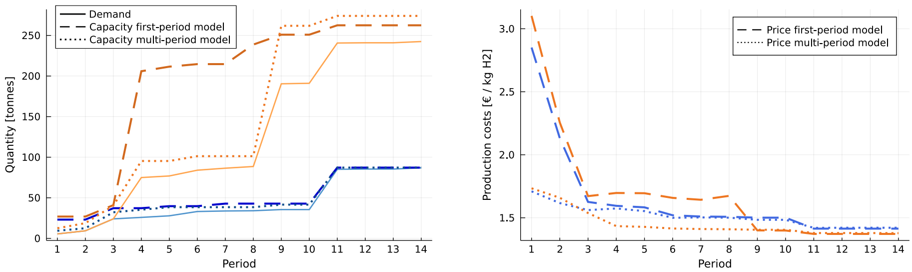
Demand scenario	maritime		all transportation	
Investment decision	first-period	multi-period	first-period	multi-period
Built facilities #	12	13	13	13
Expansion #	9	2	10	4
Total capacity [tonnes/day]	87.2	87.2	262.5	274.0
Average size [tonnes/day]	7.2	6.7	20.2	21.1
Total cost [mill. €]	606.0	578.0	1658.7	1594.1
Average hydrogen costs [€/kg]	2.73	2.61	2.53	2.43

Table 2.3: Hydrogen infrastructure characteristics.

Comparing the maritime sector and the whole transportation sector (all transportation), the installed capacity significantly increases in the scenario with higher demand, but not the number of built facilities. In the maritime sector, using the first-period model, the number of built facilities is 12. In all other cases, the number of built facilities is 13. As a result, the average facility size in the last period is almost three times higher in the scenario for the whole transportation sector comparing to the scenario for the maritime sector. The results further show that the expansion option is more often used in the first-period model as the number of expansion is 9 and 10 for the maritime and the whole transportation demand scenario, respectively. For the first-period model, expansion is the only way how increase capacity and so it leads to a higher number of expansions compared to the multi-period model which enables to build facilities later during the planning horizon.

Table 2.3 further indicates that the capacity utilization is better in the scenario for the maritime sector where the installed capacity is only slightly higher than demand in the last period. The installed capacity is 87.2 tonnes per day for both models and demanded hydrogen amount is 86.9 tonnes per day. In the whole transportation sector scenario, the infrastructure can daily provide 20 or

37 tonnes of hydrogen more than is the demanded amount for the first-period and the multi-period model, respectively.



(a) Daily demand and installed capacity

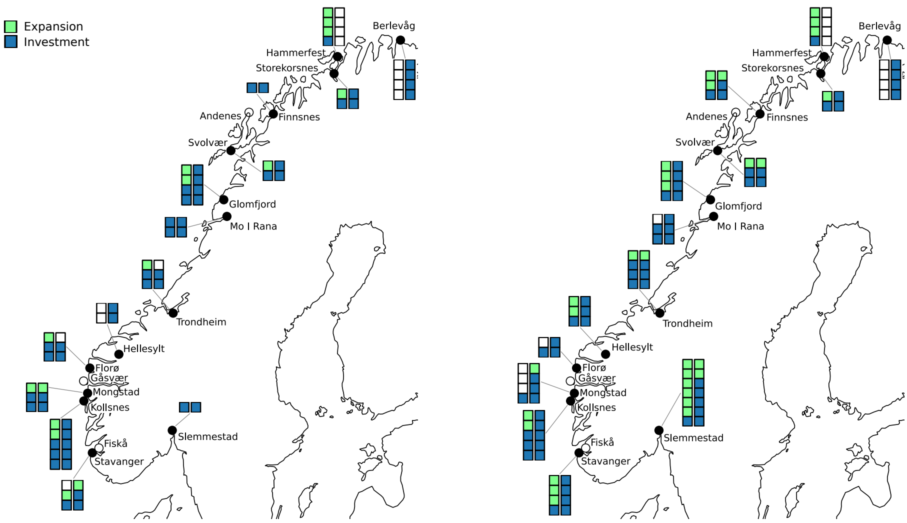
(b) Average production costs

Figure 2.4: Illustration of installed capacity and average production costs in each period for both demand scenarios and both models. Blue lines refer to the maritime scenario and orange lines to the whole transportation sector.

Figure 2.4a provides an overview of installed capacity during the planning horizon. The capacity difference between installed capacity and demand is generally low in the maritime scenario independently of the used model. In the whole transportation sector scenario, the first-period model expands in the period 4 and then the installed capacity is 2.7 times higher than the demand. In the multi-period model, the significant increase in capacity comes in period 9 where three of the four expansion in this scenario are performed and then the increase in capacity is significantly higher than the increase in demand. However, the difference is much lower than in the first-period model and the low capacity utilization affects only periods 9 and 10. The reason is that expansion is allowed only once so the expansion is performed directly to the target size. Figure 2.4a also shows that from period 11 onwards, demand remains constant and the installed capacity is just slightly higher than demand because the investment and expansion decision aimed to satisfy this target value of demand. In addition, the choice of capacities is limited by the discrete available capacities. With our choice of discrete capacities, the lower demand in the maritime scenario can be satisfied with low excess capacity. When larger capacities are needed, it becomes more difficult to successively build the capacity in line with growing demand because differences between adjacent capacities are increasing.

Figure 2.4b shows the average hydrogen production costs. In the first three periods, the multi-period model performs significantly better because it allows to build only a few facilities with high utilization in the first periods and to build more later when demand increases. This advantage of the multi-period model also leads to lower total costs and about 5% lower average hydrogen costs than the first-period model.

Figure 2.4 as a whole further illustrates the economies of scale, as increasing demand leads to lower unit production costs. We can see an exception in the first-period model in the scenario for the whole transportation sector. The average production costs in period 4 and 5 are higher than the costs in period 3 and then again the costs increases in period 8. Due to the capacity expansion in period 4 and 8 (see Figure 2.4a), the increase in capacity is significantly higher than the demand growth. The capacity utilization is low, and the unit production costs increase.



(a) Maritime transportation (b) Whole transportation sector

Figure 2.5: Investment and expansion structure of opened hydrogen facilities. The column height corresponds to the installed discrete capacity. Left columns represent the first-period model and right column represent the multi-period model.

The optimal investments in opening and expanding facilities for both models and both demand scenarios are illustrated in Figure 2.5. Figure 2.5a shows the hydrogen production infrastructure for maritime transportation and Figure 2.5b shows the infrastructure when the whole transportation sector is considered. The blue boxes denote the discrete capacity that was originally invested in, and the green boxes represent the additional discrete expansion capacity. Comparing the Figures 2.5a and 2.5b, there is no big difference in the infrastructure design in the northern part of Norway because most of the demand in that region comes from the maritime sector. The main difference and also the highest density of opened facilities is in the southern part of Norway. In the maritime scenario in the first-period model (left column), the facilities in Mongstad and Florø are larger than in the whole transportation sector scenario even if the demand in the whole transportation sector is higher and the basic demand from the maritime sector is the same. In the whole transportation scenario, the facility in Slemmestad expands already in period 4 to the target size (see Figure 2.4) and helps to satisfy the demand on the west coast.

The demand in the first periods is very low. Because of that, the infrastructure has to be successively built to satisfy demand from the first period. Later, when demand increases, there are already several smaller facilities that still have to be used and satisfy a part of this demand. The remaining requested hydrogen amount is not large enough to build a new SMR+ facility. An SMR+ facility is favourable for quantities higher than 210 tonnes hydrogen daily which is just slightly lower than hydrogen demand in the last period. As a result, due to the low initial demand level, there are built smaller EL facilities in all tested cases.

2.5 Conclusion

We study the optimal hydrogen infrastructure for maritime transportation in Norway. We use two multi-period models and analyze two demand scenarios. We consider capacitated modular facility location problem with economies of scale and two possible production technologies. We allow the production rate to differ from the installed capacity for both technologies.

Scenario with higher demand does not lead to a higher number of built facilities suggesting that the maritime sector can help to create a hydrogen infrastructure that can be used for the whole transportation sector later. Due to economies of scale, increasing demand with a stable number of facilities leads to lower production costs. This further indicates that higher initial demand could help to achieve higher competitiveness of hydrogen.

The impact of hydrogen demand generated by the road traffic sector on the size of the Slemmestad facility reflects that it would be worth considering candidate facility locations in the inland southern part of Norway.

As the investment decision flexibility has a significant impact on the designed infrastructure, a natural extension of this work is to allow facility closing and technology change during the planning horizon.

The infrastructure design and overall costs highly depend on the demand scenario. An extension of this work is to introduce uncertain demand and thus several demand scenarios and construct a stochastic optimization model. It will be also interesting to analyze the technology choice and the cost structure if we consider uncertainty in costs.

Considering international maritime transportation, ships may purchase fuel in foreign countries. It may increase the uncertainty in demand and lead to pressure on the hydrogen price in Norway. A complex model where the impact of international hydrogen purchasing on national hydrogen demand and hydrogen price is studied is subject to future work.

Bibliography

Aarskog, F. G. and Danebergs, J. (2020). Estimation of energy demand in the Norwegian high-speed passenger ferry sector towards 2030. IFE/E-2020/003, Halden, Norway.

Aarskog, F. G., Danebergs, J., Strømgren, T., and Ulleberg, Ø. (2020). Energy and cost analysis of a hydrogen driven high speed passenger ferry. *International Shipbuilding Progress*, 67(1):97–123.

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

Arabani, A. B. and Farahani, R. Z. (2012). Facility location dynamics: An overview of classifications and applications. *Computers & Industrial Engineering*, 62(1):408–420.

Behmardi, B. and Lee, S. (2008). Dynamic multi-commodity capacitated facility location problem in supply chain. In *Proceedings of the 2008 Industrial Engineering Research Conference*, pages 1914–1919. Institute of Industrial and Systems Engineers (IISE).

Castro, J., Nasini, S., and Saldanha da Gama, F. (2017). A cutting-plane approach for large-scale capacitated multi-period facility location using a specialized interior-point method. *Mathematical Programming*, 163(1-2):411–444.

Correia, I. and Captivo, M. E. (2003). A Lagrangean heuristic for a modular capacitated location problem. *Annals of Operations Research*, 122(1):141–161.

Damman, S., Sandberg, E., Rosenberg, E., Pisciella, P., and Johansen, U. (2020). Largescale hydrogen production in Norway – possible transition pathways towards 2050. *SINTEF Rapport 2020-00179, Trondheim, Norway.*

Danebergs, J. and Aarskog, F. G. (2020). Future compressed hydrogen infrastructure for the domestic maritime sector. IFE/E-2020/006, Halden, Norway.

Dias, J., Captivo, M. E., and Clímaco, J. (2007). Dynamic location problems with discrete expansion and reduction sizes of available capacities. *Investigação Operacional*, 27(2):107–130.

DNV GL (2019). Produksjon og bruk av hydrogen i Norge. Rapport 2019-0039, Oslo, Norway, (in Norwegian).

Fridstrøm, L., Tomasgard, A., Eskeland, G. S., Espegren, K. A., Rosenberg, E., Helgesen, P. I., Lind, A., Ryghaug, M., Berg, H. B., Walnum, H. J., Ellingsen, L., and Graabak, I. (2018). Decarbonization of transport, a position paper prepared by FME MoZEES and FME CenSES. ISBN 978-82-93198-25-3.

Han, J.-H., Ryu, J.-H., and Lee, I.-B. (2012). Modeling the operation of hydrogen supply networks considering facility location. *International Journal of Hydrogen Energy*, 37(6):5328–5346.

Hirth, M., Hove, K., Janzen, D., Eide, P., Helland, P., Østvik, I., Ryberg, T., and Ødegard, J. (2019). Norwegian future value chains for liquid hydrogen. NCE Maritime CleanTech, Report liquid hydrogen 2019, Stord, Norway.

Holmberg, K. (1994). Solving the staircase cost facility location problem with decomposition and piecewise linearization. *European Journal of Operational Research*, 75(1):41–61.

Jakobsen, D. and Åtland, V. (2016). Concepts for large scale hydrogen production. Master's thesis, Department of Energy and Process Engineering, NTNU, Trondheim, Norway.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2015). Dynamic facility location with generalized modular capacities. *Transportation Science*, 49(3):484–499.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2016). Solving a dynamic facility location problem with partial closing and reopening. *Computers & Operations Research*, 67:143–154.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2017). Lagrangian heuristics for large-scale dynamic facility location with generalized modular capacities. *INFORMS Journal on Computing*, 29(3):388–404.

Keipi, T., Tolvanen, H., and Konttinen, J. (2018). Economic analysis of hydrogen production by methane thermal decomposition: Comparison to competing technologies. *Energy Conversion and Management*, 159:264–273.

Luss, H. (1982). Operations research and capacity expansion problems: A survey. *Operations Research*, 30(5):907–947.

Mäkitie, T., Hanson, J., Steen, M., Hansen, T., and Andersen, A. D. (2020). The sectoral interdependencies of low-carbon innovations in sustainability transitions. FME NTRANS Working paper 01/20, Trondheim, Norway.

Melo, M. T., Nickel, S., and Saldanha da Gama, F. (2009). Facility location and supply chain management–a review. *European Journal of Operational Research*, 196(2):401–412.

Myklebust, J., Holth, C., Tøftum, L. E. S., and Tomasgard, A. (2010). Optimizing investments for hydrogen infrastructure in the transport sector. In *Techno-economic modelling of value chains based on natural gas:with consideration of CO_2 emissions*, pages 27–70. Doctoral thesis at NTNU 2010:83, Department of Industrial Economics and Technology Management, Trondheim, Norway.

NEL Hydrogen (2015). Efficient electrolysers for hydrogen production. `http://wpstatic.idium.no/www.nel-hydrogen.com/2015/03/Efficient_Electrolysers_for_Hydrogen_Production.pdf/`. last accessed 05.02.2021.

Nickel, S. and Saldanha da Gama, F. (2019). Multi-period facility location. In *Location Science*, pages 303–326. Springer.

Ocean Hyway Cluster (2020a). 2030 hydrogen demand in the Norwegian domestic maritime sector. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Ocean Hyway Cluster (2020b). Interactive map - potential maritime hydrogen in Norway. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Schütz, P. (2009). Managing uncertainty and flexibility in supply chain optimization. Doctoral thesis at NTNU 2009:89, Department of Industrial Economics and Technology Management, Trondheim, Norway.

Schütz, P., Stougie, L., and Tomasgard, A. (2008). Stochastic facility location with general long-run costs and convex short-run costs. *Computers & Operations Research*, 35(9):2988–3000.

Shulman, A. (1991). An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations Research*, 39(3):423–436.

Torres-Soto, J. E. and Üster, H. (2011). Dynamic-demand capacitated facility location problems with and without relocation. *International Journal of Production Research*, 49(13):3979–4005.

Ulleberg, Ø. and Hancke, R. (2020). Techno-economic calculations of small-scale hydrogen supply systems for zero emission transport in Norway. *International Journal of Hydrogen Energy*, 45(2):1201–1211.

Van den Broek, J., Schütz, P., Stougie, L., and Tomasgard, A. (2006). Location of slaughterhouses under economies of scale. *European Journal of Operational Research*, 175(2):740–750.

Paper II

Multi-period facility location and capacity expansion with modular capacities and convex short-term costs

Šárka Štádlerová, Peter Schütz, Asgeir Tomasgard

This paper is submitted for publication and is therefore not included.

Submitted to an international journal

Paper III

Locating hydrogen production in Norway under uncertainty

Šárka Štádlerová, Trygve Magnus Aglen, Andreas Hofstad, Peter Schütz

Published in Lecture Notes in Computer Science

Paper 3
Locating hydrogen production in Norway under uncertainty

Abstract

In this paper, we study a two-stage stochastic multi-period facility location and capacity expansion problem. The problem is motivated by the real-world problem of locating facilities for green hydrogen in Norway. We formulate a model with modular capacities. Investment in a facility and expansion costs represents long-term costs. For each capacity, we define a convex short-term production cost function which enables to capture economies of scale in investment as well as in production. The objective is to minimize the total expected investment, expansion, production and distribution costs while satisfying demand in each scenario. We solve the problem using sample average approximation. The results from solving the problem show that the stochastic problem leads to lower installed capacity in the opening decisions than the expected value problem.

Keywords: Stochastic Facility Location, Capacity Expansion, Hydrogen supply chain

4.1 Introduction

In February 2020, Norway adopted more ambitious emission reduction targets than agreed upon in the Paris Agreement. The new target is to reduce greenhouse gas (GHG) emissions by at least 50% towards 2030, compared to the 1990 level (Regjeringen, 2021). To achieve this goal, the emissions from the transport sector also need to be halved. With a share of more than 30%, the transportation sector is an important contributor to total GHG emissions (Samferdselsdepartementet, 2021).

One of the key instruments for achieving the emission reduction targets is to use green hydrogen as a zero-emission energy carrier (Samferdselsdepartementet, 2021). Only hydrogen coming from a CO_2-free production process can be considered a green zero-emission fuel. Electrolysis (EL) using energy from renewable sources is the most mature technology for green hydrogen production (IRENA, 2020). EL is a quite flexible production technology and can produce in a range of 20 – 100% of installed capacity (NEL Hydrogen, 2015). The production costs are subject to economies of scale as higher production quantities result in lower average unit costs (Hirth et al., 2019).

In order to start the transition towards hydrogen in Norway, municipalities can require the usage of hydrogen as fuel when public transport contracts for ferries, high-speed passenger vessels, and coastal routes are renewed. Hydrogen is also a promising energy carrier for long-distance buses and heavy trucks (DNV GL, 2019). The Norwegian government is also working on designing possible low- and zero-emission requirements for offshore supply vessels (Nærings- og skeridepartementet, 2020). The conversion potential to zero-emission energy carrier of the offshore fleet with respect to the fleet composition and future demand is presented in Ocean Hyway Cluster (2020a). Future hydrogen demand is highly uncertain because the market share of hydrogen vehicles in the road traffic sector and the future energy carrier in the offshore sector are also subject to uncertainty.

In this paper, we study the problem of locating hydrogen production facilities in Norway under uncertain demand. We formulate our problem as a two-stage stochastic multi-period facility location problem with capacity expansion. We consider modular capacities in order to model economies of scale. The goal is to minimize expected investment, expansion, production and distribution costs of satisfying the customer demand. We distinguish between long-term investment costs and short-term operational costs to capture economies of scale in investment and production. This approach also enables the modelling of different utilization of the installed capacity. The problem is solved using sample average approximation (SAA). We compare the first-stage solution of the stochastic problem (SP) and the expected value problem (EVP) and discuss the value of the stochastic solution. We analyse the hydrogen production infrastructure and provide a managerial insight into the investment capacity of new facilities.

The remainder of this paper is structured as follows: we first provide an overview of related work to deterministic and stochastic facility location and capacity expansion problems in Section 4.2. We formulate the mathematical model for the stochastic two-stage multi-period facility location problem in Section 4.3. The solution approach is presented in Section 4.4. Case study and Computational results are discussed in Section 4.5 and 4.6, respectively. We conclude in Section 4.7.

4.2 Related work

We structure the related work into three main parts. First, we focus on literature related to deterministic facility location and capacity expansion problems before we continue with two-stage facility location and supply chain design problems. Finally, we present literature related to SAA.

Deterministic multi-period facility location and capacity expansion problems with modular capacities are studied in Shulman (1991), Dias et al. (2007). In these papers, both capacity expansion and capacity reduction are allowed. Expansion is modelled as new-building of another facility at a given location while capacity reduction means closing some or all of the facilities at a given location. An approach where capacity expansion is modelled as a modification of an existing facility is presented in Jena et al. (2015, 2016, 2017), Štádlerová and Schütz (2021). In Štádlerová and Schütz (2021), the number of capacity expansions is limited, and capacity reduction is not allowed. In Jena et al. (2015), capacity expansion and reduction are allowed multiple times. An extended version of the model from Jena et al. (2015) for multiple commodities is presented in Jena et al. (2016, 2017). See also the review Melo et al. (2009), Nickel and Saldanha da Gama (2019) for an overview over multi-period facility location problems.

Uncertainty in demand in two-stage stochastic problems is more commonly found in single-period facility location problems. The first-stage decisions usually refer to the opening of facilities and determining their capacities, while the second-stage decisions are related to distribution and demand satisfaction. A model with random demand and non-linear cost function to model economies of scale is discussed in Balachandran and Jain (1976), Schütz et al. (2008). The problem in Schütz et al. (2008) is solved using Lagrangian relaxation.

A two-stage facility location problem with depots is presented in Litvinchev and Ozuna Espinosa (2012) and also solved by Lagrangian relaxation. The model presented in Litvinchev and Ozuna Espinosa (2012) can be solved by an effective genetic algorithm as shown in Fernandes et al. (2014). A two-stage multi-period facility location model with a capacity expansion is studied in Correia and Melo (2021). The authors compare two model formulations: In the first model, capacity expansion is a part of the first-stage decisions while in the second model, capacity expansion is a second-stage decision. A multi-stage formulation of a multi-period stochastic problem is discussed in Ahmed et al. (2003).

Supply chain network design problems are similar to facility location problems and have received lots of attention. A study on designing the hydrogen supply chain under uncertain demand with a similar decision structure to Balachandran and Jain (1976), Schütz et al. (2008) is presented in Kim et al. (2008), Nunes et al. (2015). The first-stage decisions correspond to investing in production and storage capacity during the planning horizon while the second-stage decisions correspond to the distribution plan. A two-stage stochastic programming model for minimizing the total daily costs of the hydrogen supply chain with uncertain demand is presented in Dayhim et al. (2014). Compared to previous work in the hydrogen supply chain, the authors provide emission, energy consumption and risk costs. An early literature review on dynamic facility location and supply chain problems with stochastic data can be found in Owen and Daskin (1998). A review on facility location problems under uncertainty is provided in Snyder (2006) and a recent summary on facility location problems under uncertainty is presented in Govindan et al. (2017), Correia and Saldanha da Gama (2019).

The SAA algorithm allows for solving large two-stage stochastic problems with a binary first stage. See Mak et al. (1999) and Kleywegt et al. (2001) for the details on methodology. The application of SAA to a facility location problem where the availability of opened facilities is uncertain is presented in Gade and Pohl (2009). A similar problem with facility disruptions is discussed in Li and Zhang (2018). The authors combine SAA with a scenario decomposition algorithm to solve the problem. A combined solution approach of SAA and Benders decomposition for a supply chain design problem with uncertain demand is studied in Santoso et al. (2005). A supply chain design problem with a model that captures short-term as well as long-term demand uncertainty is discussed in Schütz et al. (2009). In order to increase the number of scenarios in the sample, SAA combined with dual decomposition is applied to solve the problem.

4.3 The mathematical programming model

We study a stochastic two-stage multi-period facility location and capacity expansion problem with uncertain demand. The objective is to minimize the total expected costs.

4.3.1 Problem description

We formulate our problem as a two-stage stochastic multi-period facility location and capacity expansion problem. The goal is to minimize the sum of expected discounted investment, expansion, production and distribution costs while satisfying demand in each scenario. The decisions when and where to open and which capacity to invest in are taken before the uncertainty is disclosed. In the second stage, decisions covering capacity expansion, production, and distribution are taken. Capacity expansion is allowed only once in each scenario and only in the sense of increasing the capacity level. Once a facility is opened, it cannot be closed.

We consider a set of candidate locations and a set of customers. For each facility-customer combination, we have specific unit distribution costs. However, not all customers can be served from all facilities. The investment costs are given by the installed capacity while the production costs depend both on installed capacity and production quantity. Note that investment and production costs can depend on location. The production quantities can vary from the installed capacity. However, there is a lower and an upper limit. The lower limit is given by the minimum production quantities for each capacity. The installed capacity represents the upper limit for production. This upper limit can be extended by expansion.

We model the investment and capacity decision as a discrete choice from a set of modular capacities. Expansion is then modelled as a jump between available capacities. We consider opening a small facility and expanding it as a more expensive alternative to opening a large facility right away. These extra costs are modelled as a one-time payment when expanding. However, the short-term production costs are independent of whether the capacity results from expansion or from opening the facility right away.

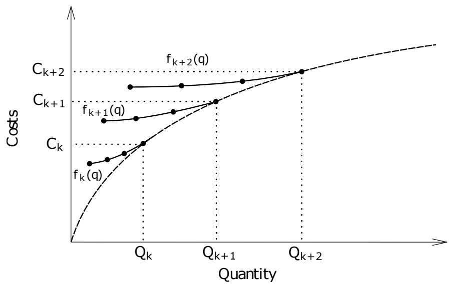


Figure 4.1: Long-term and short-term production costs

For each available capacity, we provide a piecewise linear convex short-term production cost function which enables a variation in production quantities. This approach enables to capture the economies of scale in investment as well as in production. Figure 4.1 shows our long-term (dashed line) and short-term (solid line) production costs. The capacity index of installed modular capacity is denoted k and Q_k is the appropriate quantity. The total costs for production at installed capacity k are denoted C_k. For each capacity k, we define a short-term production costs function $f_k(q)$ that enables production in a range between minimum and maximum limit. However, higher utilization of installed capacity leads to lower unit costs. This approach of modelling investment and production costs is similar to the one in Schütz et al. (2008).

4.3.2 Mathematical formulation

Let us first introduce the following notation:

Sets

$\mathscr{B}$ Set of breakpoints of the short-term cost function
$\mathscr{I}$ Set of possible facility locations
$\mathscr{J}$ Set of customer ports
$\mathscr{K}$ Set of available discrete capacities
$\mathscr{S}$ Set of scenarios
$\mathscr{T}$ Set of time periods
$\mathscr{T}_1$ Set of time periods corresponding to the first-stage, $\mathscr{T}_1 \subset \mathscr{T}$

Parameters and coefficients

C_{ik} investment costs at location i for capacity point k;
D^s_{jt} demand at customer j in period t, and scenario s;
E_{ikl} costs of expansion at location i from capacity in point k to capacity in point l;
F_{ibk} costs at location i at breakpoint b of the short-term cost function of capacity k;
L_{ij} 1 if demand at location j can be served from facility i, 0 otherwise;
Q_{bk} production volume at breakpoint b of the short-term cost function, for capacity point k;
T_{ij} distribution costs from facility i to customer j;
y_{ikl0} initial facility variable;
δ_t discount factor in period t;
p^s probability of scenario s;

Decision variables

x^s_{ijt} amount of customer demand at customer location j satisfied from facility i in period t in scenarios s;
y^s_{iklt} 1 if facility is operated in location i in period t, with originally installed capacity k, and operated capacity l in scenario s, 0 otherwise;
μ^s_{bilt} weight of breakpoint b at location i for capacity point k in period t and scenario s.

We present a two-stage stochastic multi-period model. The model is given as:

$$\min \sum_{s\in\mathscr{S}} p^s \Bigg[\sum_{i\in\mathscr{I}} \sum_{k\in\mathscr{K}} \sum_{t\in\mathscr{T}} \delta_t C_{ik} \left(y^s_{ikkt} - y^s_{ikk(t-1)} \right) + \sum_{i\in\mathscr{I}} \sum_{k\in\mathscr{K}} \sum_{l\in\mathscr{K}:l>k} \sum_{t\in\mathscr{T}} \delta_t E_{ikl} \left(y^s_{iklt} - y^s_{ikl(t-1)} \right) + \sum_{b\in\mathscr{B}} \sum_{i\in\mathscr{I}} \sum_{l\in\mathscr{K}} \sum_{t\in\mathscr{T}} \delta_t F_{ibl} \mu^s_{bilt} + \sum_{i\in\mathscr{I}} \sum_{j\in\mathscr{J}} \sum_{t\in\mathscr{T}} \delta_t T_{ij} x^s_{ijt} \Bigg], \tag{4.1}$$

subject to:

$$\sum_{k\in\mathscr{K}} \sum_{l\in\mathscr{K}:l\geq k} y^s_{iklt} \leq 1, \qquad i\in\mathscr{I}, t\in\mathscr{T}, s\in\mathscr{S}, \tag{4.2}$$

$$\sum_{k\in\mathscr{K}} \sum_{l\in\mathscr{K}:l>k} y^s_{iklt} = 0, \qquad i\in\mathscr{I}, t\in\mathscr{T}_1, s\in\mathscr{S}, \tag{4.3}$$

$$\sum_{t'=1}^{t-1} y^s_{ikkt'} \geq \sum_{l\in\mathscr{K}:l>k} y^s_{iklt}, \qquad i\in\mathscr{I}, k\in\mathscr{K}, t\in\mathscr{T}, s\in\mathscr{S}, \tag{4.4}$$

$$\sum_{l\in\mathscr{K}:l\geq k} y^s_{iklt} \geq \sum_{l\in\mathscr{K}:l\geq k} y^s_{ikl(t-1)}, \qquad i\in\mathscr{I}, k\in\mathscr{K}, t\in\mathscr{T}, s\in\mathscr{S} \tag{4.5}$$

$$y^s_{iklt} - y^s_{ikl(t-1)} \geq 0, \qquad i\in\mathscr{I}, k\in\mathscr{K}, l\in\mathcal{K}: l>k, t\in\mathscr{T}, s\in\mathscr{S}, \tag{4.6}$$

$$\sum_{b\in\mathscr{B}} \mu^s_{bilt} = \sum_{k\in\mathscr{K}} y^s_{iklt}, \qquad i\in\mathscr{I}, l\in\mathscr{K}, t\in\mathscr{T}, s\in\mathscr{S}, \tag{4.7}$$

$$\sum_{j\in\mathscr{J}} x^s_{ijt} = \sum_{b\in\mathscr{B}} \sum_{l\in\mathscr{K}} Q_{bl} \mu^s_{bilt}, \qquad i\in\mathscr{I}, t\in\mathscr{T}, s\in\mathscr{S}, \tag{4.8}$$

$$\sum_{i\in\mathscr{I}} x^s_{ijt} = D^s_{jt}, \qquad j\in\mathscr{J}, t\in\mathscr{T}, s\in\mathscr{S}, \tag{4.9}$$

$$x^s_{ijp} \leq L_{ij} D^s_{jt}, \qquad i \in \mathscr{I}, j \in \mathscr{J}, t \in \mathscr{T}, s \in \mathscr{S}, \quad (4.10)$$

$$\frac{1}{|\mathscr{S}|} \sum_{s' \in \mathscr{S}} \sum_{l \in \mathscr{K}: l \geq k} y^{s'}_{iklt} = \sum_{l \in \mathscr{K}: l \geq k} y^s_{iklt}, i \in \mathscr{I}, k \in \mathscr{K}, t \in \mathscr{T}, s \in \mathscr{S}, \quad (4.11)$$

$$y^s_{iklt} \in \{0, 1\}, \qquad i \in \mathscr{I}, k \in \mathscr{K}, l \in \mathscr{K} : l \geq k, t \in \mathscr{T}, s \in \mathscr{S}, \quad (4.12)$$

$$x^s_{ijt} \geq 0, \qquad i \in \mathscr{I}, j \in \mathscr{J}, t \in \mathscr{T}, s \in \mathscr{S} \quad (4.13)$$

$$\mu^s_{bilt} \geq 0, \qquad b \in \mathscr{B}, i \in \mathscr{I}, k \in \mathscr{K}, t \in \mathscr{T}, s \in \mathscr{S}. \quad (4.14)$$

The objective function (4.1) is equal to the expected discounted costs of investment, expansion, production and distribution costs. Restrictions (4.2) guarantee that only one facility is opened at a given location and that this facility is operated at only one capacity at a time. Constraints (4.3) ensure that we are allowed to open facilities in the first stage, but not to expand them. Restrictions (4.4) make sure that only previously opened facilities can be expanded and constraints (4.5) ensure that a facility can be expanded but cannot be closed. Capacity expansion is allowed only once during the planning horizon in each scenario. The variable y^s_{iklt} contains information about the initially installed capacity k as well as the capacity l at which it is currently operated. After expansion, the operated capacity l is higher than the installed capacity k. Inequalities (4.6) ensure that capacity index l can change only once. Equations (4.7) guarantee that production is allocated only to opened facilities and that the short-term production cost function depends on the operated capacity. Equations (4.8) express the requirement that the whole production has to be distributed to customers. Equations (4.9) ensure demand satisfaction in each scenario, while constraints (4.10) specify if customer j can be served from facility i.

Constraints (4.11) are the non-anticipativity constraints (see e.g. Rockafellar and Wets (1991)) that ensure that the opening capacity k is the same in all scenarios. Once a facility has been opened with capacity k in a given scenario s, it has to be operated at a capacity $l \geq k$. Hence, the right-hand side, $\sum_{l \in \mathscr{K}: l \geq k} y^s_{iklt}$, is equal to 1. The left-hand side then ensures that the facility is opened with capacity k in all scenarios, even though it might be operated at different capacities l in different scenarios.

Restrictions (4.12)–(4.14) are the binary and non-negativity requirements for the decision variables. The variables are defined for each scenario. However, investment decisions must be taken before the uncertainty is disclosed.

4.4 Solution approach

We use the SAA algorithm (Mak et al., 1999), (Kleywegt et al., 2001) to solve our two-stage stochastic multi-period model with binary variables. A description of the algorithm can also be found in Santoso et al. (2005) and Schütz et al. (2009), but we summarize it here for the sake of completeness. Using the SAA approach, the problem is repeatedly solved with a smaller set of scenarios. First, a random sample $\xi^1, ..., \xi^n$ with a size N is generated. Then the expectation $\mathbb{E}[Q(y,\xi)]$ is approximated by the sample average function $\frac{1}{N}\sum_{n=1}^{N} Q(y,\xi^n)$. We approximate our problem with the following SAA problem:

$$\min \left\{ \hat{g}(y) = c^T y + \frac{1}{N}\sum_{n=1}^{N} Q(y,\xi^n) \right\} \tag{4.15}$$

With increasing sample size, the optimal solution of (4.15), $\hat{y}_N$ converges to the optimal solution of the original problem with probability one. In practical implementations, the sample size is often chosen with respect to the computational effort. As we have issues solving our model with more than 10 scenarios, we follow the approach from Santoso et al. (2005). The authors show that a higher number of samples can be more efficient than increasing the number of scenarios.

Let M be the number of independent samples and v_N^m the optimal objective function of a problem for $m = 1, ..., M$. The average objective function value is then computed as:

$$\overline{v}_{N,M} = \frac{1}{M}\sum_{m=1}^{M} v_N^m \tag{4.16}$$

Equation (4.16) represents a statistical lower bound (LB) on the objective function value for the original problem (Mak et al., 1999), (Norkin et al., 1998).

Let $N' >> N$ be the reference sample representing the true uncertainty in the problem and $\overline{y}$ a feasible first-stage solution. Then, the objective function of the original problem for a given solution $\overline{y}$ can be calculated as:

$$\widetilde{g}_{N'}(\overline{y}) = c^T\overline{y} + \frac{1}{N'}\sum_{n=1}^{N'} Q(\overline{y},\xi^n) \tag{4.17}$$

Equation (4.17) provides an upper bound (UB) on the optimal objective function value. Having the lower and upper bound estimates, we can compute the estimated optimality gap as:

$$gap_{N,M,N'}(\overline{y}) = \widetilde{g}_{N'}(\overline{y}) - \overline{v}_N^m. \tag{4.18}$$

4.5 Case study

In this section, we provide the real-world input data used for solving the problem of locating hydrogen production in Norway under uncertainty.

4.5.1 Facilities and production

We consider 17 candidate locations for the opening of new facilities on the Norwegian west coast. The candidate locations are taken from Ocean Hyway Cluster (2020b). We approximate the facility capacity by 8 discrete points and provide the investment and production costs at full capacity utilization for EL in Table 4.1.

Discrete capacity	1	2	3	4	5	6	7	8
Capacity [tonnes/day]	0.6	3.1	6.2	12.2	30.3	61.0	151.5	304.9
Investment EL [mill. €]	1.4	6.0	11.2	20.5	46.5	87.2	197.7	371.5
Production EL [€/kg]	1.95	1.61	1.53	1.45	1.43	1.42	1.40	1.38

Table 4.1: Investment and production costs at full capacity utilization for EL (Štádlerová and Schütz, 2021)

There are minimum production requirements for electrolysis, as the production rate can decrease towards 20% of the installed capacity. We approximate the short-term production costs by a convex piecewise linear function with three linepieces. We define four breakpoints at $20\%, 50\%, 80\%$, and 100% of installed production quantity. The 20% breakpoint represents the minimum production requirement based on the technical specifications for electrolysis, and the 100% breakpoint represents full utilization of installed capacity. Each breakpoint is characterized by a specific production quantity and production costs. We can produce arbitrary quantities from the range between $20 - 100\%$ of the installed capacity by a linear combination of two neighbourhood breakpoints. The short-term costs at a breakpoint are calculated based the a model provided in Jakobsen and Åtland (2016). We assume that the investment and production costs are independent of facility location.

We calculate the expansion costs E_{ikl} as: $E_{ikl} = (C_{il} - C_{ik}) \cdot (100 + \alpha)\%$. The expansion costs are equal to the difference between investment costs of opening a facility with capacity l and a facility with capacity k, where $k < l$, plus an additional mark-up α. In our case, the mark-up α is 10%

We use the distribution costs for compressed hydrogen provided by Danebergs and Aarskog (2020). We consider demand points that aggregate customer demand from the whole municipality, and if a demand point is located in the same municipality as a facility, we assume zero distribution costs. The reason is that

Distance [km]	1-50	51-100	101-200	201-400	401-800	801-1000
Costs	0.00498	0.00426	0.00390	0.00372	0.00363	0.00360

Table 4.2: Hydrogen distribution costs in [€/km/kg H_2] (Danebergs and Aarskog, 2020)

the starting point for our case study is the production of hydrogen for maritime transportation. The demand points for this sector are limited to ports. For locations along the Norwegian coastline, we assume that hydrogen production will take place in port or close to the port with negligible distribution costs. This assumption has then been extended to municipalities producing hydrogen for other sectors than maritime for reasons of consistency. We set the distance limit between a production facility and a customer to 1000 km. See Table 4.2 for the distribution costs for compressed hydrogen. The production cost and distribution cost data for our case are identical to the from Šťádlerová and Schütz (2021). For simplification, we assume that the discount factor is equal to one in each period.

4.5.2 Demand

We consider three main demand components. In the maritime sector, the hydrogen demand estimations are based on current ferry routes and the assumption that the new public contracts will require hydrogen as an energy carrier (Ocean Hyway Cluster, 2020a), (Aarskog and Danebergs, 2020). The demand estimations in the land-based sector from DNV GL (2019) are based on the emission reduction goal within 2030 stated in Samferdselsdepartementet (2021). In the offshore sector, we use the hydrogen demand estimations from Aglen and Hofstad (2022). These estimations are based on the medium penetration scenario from Ocean Hyway Cluster (2020a) which calculates the energy consumption for ammonia. However, hydrogen fuel alternative is just as likely to occur (Ulstein Design, 2021). These different demand components are shown in Figure 4.2 together with the expected demand level and the maximum potential hydrogen demand consisting of all three components. The maritime demand is quite certain. Thus, it represents the minimum demand level and is present in all demand scenarios.

We aggregate individual customer demand into 70 demand points located in Norway. These demand points consist of 51 ports that are relevant for the maritime and the offshore sector and 19 municipalities with the highest road traffic volumes according to the statistic provided in Statistics Norway (2018). Based on the traffic volumes statistic (Statistics Norway, 2018), we divide the road traffic demand among the different municipalities. We remove municipalities with demand lower than 3.65 tonnes H_2/year. However, not all customers, respectively

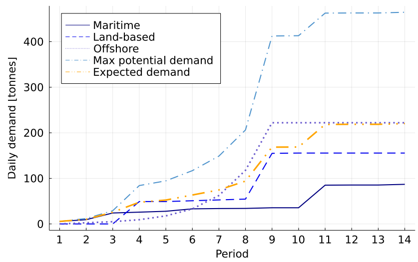


Figure 4.2: Demand development

demand points, have demand in all scenarios.

Our planning horizon is 14 periods. Demand is non-decreasing during the whole planning horizon in all considered sectors. In the maritime sector, demand is slightly increasing until period 10 and there is a jump in period 11 when the coastal route Bergen-Kirkenes is to be operated on hydrogen fuels. The jumps in the land-based sector correspond to the strategic government plan to start with the transition towards hydrogen for buses and trucks. The offshore sector will not start the transition towards hydrogen before period 4.

The market share of hydrogen vehicles and hydrogen-driven offshore supply vessels is highly uncertain. We consider demand in the land-based sector and offshore sector to represent a conversion potential and assume that the probability of reaching the maximal potential demand is low. Therefore, we assume that our demand scenarios are not evenly distributed between the minimum and the maximum potential demand. We assume the expected value to be a weighted average of minimum and maximum demand with coefficients 0.65 and 0.35, respectively.

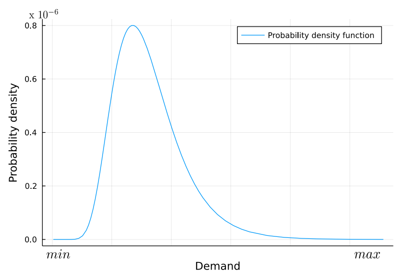


Figure 4.3: Probability density function for hydrogen demand

We expect that scenarios with lower demand consisting of maritime demand and a share of the land-based and offshore sector are more likely to occur than very optimistic hydrogen scenarios with very high demand. Thus, we need a left-skewed distribution with a low probability of extreme values to sample the scenarios from. We therefore assume a log-normal distribution, $D \sim Lognormal(\mu, \sigma^2)$. The expected value $E(D)$ is given by the previously computed expected demand level and we assume the standard deviation to be $\sigma = 0.3$ as this value still allows some of the high demand scenarios to occur. The probability density function of our log-normal distribution is shown in Figure 4.3.

4.6 Computational results

The model is implemented in Julia 1.6.5 and solved using Gurobi Optimizer version 9.5. All calculations have been run on a computer with two 3.6 GHz Intel Xeon Gold 6244 CPU (8 core) processors and 384 GB RAM.

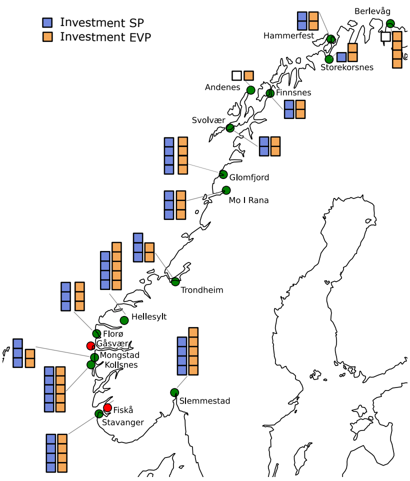
The problem (4.15) is solved for $M = 50$ SAA problems where each of the problems has a sample size of $N = 10$. The reference sample size is $N' = 1000$ and we evaluate the performance on the reference sample for each of the 50 SAA solutions. We choose to solve the problems (4.15) with relative optimality gap $\gamma' < 2\%$.

Problem	LB [x10^6 €]	UB [x10^6 €]	$gap_{N,M,N'}(\overline{y})[\%]$
SP	1381.2	1455.2	5.36
EEV	-	∞	-

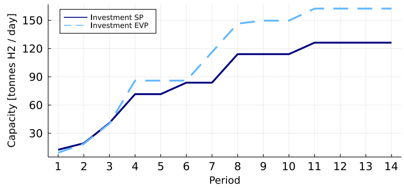
Table 4.3: Evaluation of the SP and the EEV

We show the best statistical lower and upper bound of the SP in Table 4.3 and compare the results with the EVP. We calculate the expected value of the EVP solution (EEV) and compare the results with the SP. The value of the stochastic solution is: $VSS = EEV - SP$ Birge and Louveaux (2011). The results show that the EVP solution is infeasible. Thus, the VSS goes to infinity. This shows that even if the EVP problem is easier to solve and we can find an optimal solution, it is important to consider the uncertainty in our problem.

To analyze the first stage decisions, we study the opening decisions in the SP and the EVP. Figure 4.4a illustrates the facility locations and the opening size of facilities before expansion. When comparing the number of opened facilities, we open 13 facilities in the SP and 15 facilities in the EVP. However, in general, the differences between the SP and the EVP are very small. The main differences can be seen in the northern part of Norway where we do not open a facility in Berlevåg and Andenes in the SP. Thus, we install more capacity in the EVP in comparison to SP. However, the infeasibility comes from the south-western part of Norway even if the number of opened facilities is equal. Please note that the difference between capacity 2 and 3 is only 3.1 tonnes daily while the difference



(a) Location and opening capacity of hydrogen facilities in the SP and the EVP



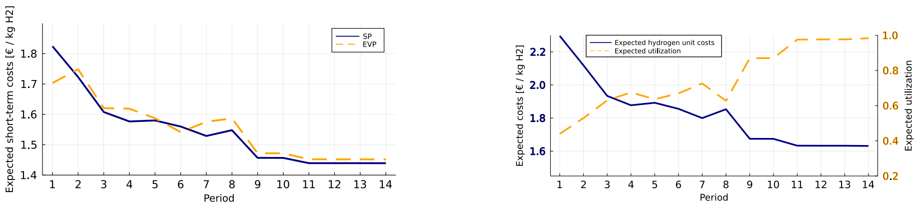
(b) Opening capacity in the SP and the EVP

Figure 4.4: First-stage decisions: Investment in the SP and the EVP

between capacity 4 and 5 is 18.1 tonnes daily. Thus, we install more capacity in the EVP as we open two large facilities in Hellesylt and Slemmestad. Most of the land-based demand is located in the south-western part of Norway and this area is also affected a lot by the offshore demand. Thus, here, we observe the highest differences between the scenarios and the large capacities installed for the EVP cause infeasibility for scenarios with low demand. In the EVP, we cannot fulfil the minimum production requirements for scenarios with low demand due to the large facilities in Hellesylt and Slemmestad.

The development of installed capacity in the first stage in the SP and the EVP solution can be seen in Figure 4.4b. The installed capacity is almost the same in the first three periods because the differences between scenarios are low until period three. Then, both lines indicate growing capacity. However, the installed capacity in SP is considerably lower. The solution of the SP leads to more conservative investment decisions and additional capacity is installed in the expansion step. The expected demand level is considerably higher than the minimum demand so the EVP problem leads to more extensive investments than the SP which is also the reason for the infeasibility of the EVP.

For illustration, we show the expected unit short-term costs in the SP and the EVP in Figure 4.5a. Please note that we show results for a feasible subset



(a) Expected unit short-term costs in the SP and the EVP

(b) Expected unit costs and expected capacity utilization in the SP

Figure 4.5: Expected hydrogen costs

of scenarios in the EVP. The EVP provides lower costs in the first period due to the lower installed capacity (see Figure 4.4b) resulting in higher utilization. In the following periods, the costs in the SP are, in general, lower. However, the costs are very similar because expansion in the second stage provides a lot of flexibility to adjust the infrastructure as a reaction to growing demand. Expected unit hydrogen costs and expected utilization for the SP are shown in Figure 4.5b. The expected unit hydrogen costs have a decreasing tendency that is in line with the growing capacity (see Figure 4.4b) and increasing utilization. The unit production costs have two peaks in period 5 and 8 that are related to a decrease in capacity utilization as lower utilization results in higher unit costs. In expectation, unit production costs are decreasing together with increasing capacity and its utilization which indicates the presence of economies of scale in hydrogen production.

4.7 Conclusion

We study the optimal hydrogen production infrastructure under uncertain demand in Norway. We present a model for a two-stage stochastic multi-period facility location problem with capacity expansion. The problem is hard to solve and using commercial software, we can solve it with 10 scenarios. Therefore, we use SAA to solve the problem. This approach provides good solutions with an estimated gap between the lower and the upper bound of 5.36%.

The quality of the solution is limited by the number of scenarios we can solve the problem with. Implementing an efficient solution method in order solve the problem with more scenarios and thus improve the solution quality is a natural extension of this work.

Another extension of this work is to study how the investment structure will change when we modify the underlying demand distribution.

Expansion in the second stage provides a lot of flexibility in terms of reaction to growing demand. It is worth considering, how the investment decisions will change for different models. We can consider a multi-stage model, or a more restrictive model where expansion is the first-stage decision and only decisions regarding demand allocation are taken in the second stage. In future work, uncertainty in costs might be considered as well.

Bibliography

Aarskog, F. G. and Danebergs, J. (2020). Estimation of energy demand in the Norwegian high-speed passenger ferry sector towards 2030. IFE/E-2020/003, Halden, Norway.

Aglen, T. M. and Hofstad, A. (2022). Designing the hydrogen supply chain for maritime transportation. Master's thesis, Department of Industrial Economics and Technology Management, NTNU, Trondheim, Norway.

Ahmed, S., King, A. J., and Parija, G. (2003). A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization*, 26(1):3–24.

Balachandran, V. and Jain, S. (1976). Optimal facility location under random demand with general cost structure. *Naval Research Logistics Quarterly*, 23(3):421–436.

Birge, J. R. and Louveaux, F. (2011). *Introduction to stochastic programming.* Springer Science & Business Media, New York, 2 edition.

Correia, I. and Melo, T. (2021). Integrated facility location and capacity planning under uncertainty. *Computational and Applied Mathematics*, 40(5):1–36.

Correia, I. and Saldanha da Gama, F. (2019). Facility location under uncertainty. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location science*, pages 185–213. Springer.

Danebergs, J. and Aarskog, F. G. (2020). Future compressed hydrogen infrastructure for the domestic maritime sector. IFE/E-2020/006, Halden, Norway.

Dayhim, M., Jafari, M. A., and Mazurek, M. (2014). Planning sustainable hydrogen supply chain infrastructure with uncertain demand. *International Journal of Hydrogen Energy*, 39(13):6789–6801.

Dias, J., Captivo, M. E., and Clímaco, J. (2007). Dynamic location problems with discrete expansion and reduction sizes of available capacities. *Investigação Operacional*, 27(2):107–130.

DNV GL (2019). Produksjon og bruk av hydrogen i Norge. Rapport 2019-0039, Oslo, Norway, (in Norwegian).

Fernandes, D. R., Rocha, C., Aloise, D., Ribeiro, G. M., Santos, E. M., and Silva, A. (2014). A simple and effective genetic algorithm for the two-stage capacitated facility location problem. *Computers & Industrial Engineering*, 75:200–208.

Gade, D. and Pohl, E. (2009). Sample average approximation applied to the capacitated-facilities location problem with unreliable facilities. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 223(4):259–269.

Govindan, K., Fattahi, M., and Keyvanshokooh, E. (2017). Supply chain network design under uncertainty: A comprehensive review and future research directions. *European Journal of Operational Research*, 263(1):108–141.

Hirth, M., Hove, K., Janzen, D., Eide, P., Helland, P., Østvik, I., Ryberg, T., and Ødegard, J. (2019). Norwegian future value chains for liquid hydrogen. NCE Maritime CleanTech, Report liquid hydrogen 2019, Stord, Norway.

IRENA (2020). Green hydrogen: A guide to policy making. *International Renewable Energy Agency.*

Jakobsen, D. and Åtland, V. (2016). Concepts for large scale hydrogen production. Master's thesis, Department of Energy and Process Engineering, NTNU, Trondheim, Norway.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2015). Dynamic facility location with generalized modular capacities. *Transportation Science*, 49(3):484–499.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2016). Solving a dynamic facility location problem with partial closing and reopening. *Computers & Operations Research*, 67:143–154.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2017). Lagrangian heuristics for large-scale dynamic facility location with generalized modular capacities. *INFORMS Journal on Computing*, 29(3):388–404.

Kim, J., Lee, Y., and Moon, I. (2008). Optimization of a hydrogen supply chain under demand uncertainty. *International Journal of Hydrogen Energy*, 33(18):4715–4729.

Kleywegt, A. J., Shapiro, A., and Homem-de Mello, T. (2001). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

Li, X. and Zhang, K. (2018). A sample average approximation approach for supply chain network design with facility disruptions. *Computers & Industrial Engineering*, 126:243–251.

Litvinchev, I. and Ozuna Espinosa, E. L. (2012). Solving the two-stage capacitated facility location problem by the lagrangian heuristic. In Hu, H., Shi, X., Stahlbock, R., and Voß, S., editors, *International Conference on Computational Logistics*, pages 92–103. Springer, Berlin, Heidelberg.

Mak, W.-K., Morton, D. P., and Wood, R. K. (1999). Monte carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24(1-2):47–56.

Melo, M. T., Nickel, S., and Saldanha da Gama, F. (2009). Facility location and supply chain management–a review. *European Journal of Operational Research*, 196(2):401–412.

NEL Hydrogen (2015). Efficient electrolysers for hydrogen production. http://wpstatic.idium.no/www.nel-hydrogen.com/2015/03/Efficient_Electrolysers_for_Hydrogen_Production.pdf/. last accessed 05.02.2021.

Nickel, S. and Saldanha da Gama, F. (2019). Multi-period facility location. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, pages 303–326. Springer, Cham.

Norkin, V. I., Pflug, G. C., and Ruszczyński, A. (1998). A branch and bound method for stochastic global optimization. *Mathematical programming*, 83(1):425–450.

Nunes, P., Oliveira, F., Hamacher, S., and Almansoori, A. (2015). Design of a hydrogen supply chain with uncertainty. *International Journal of Hydrogen Energy*, 40(46):16408–16418.

Nærings- og skeridepartementet (2020). Grønnere og smartere morgendagens maritime næring. https://www.regjeringen.no/no/dokumenter/meld.-st.-10-20202021/id2788786/. last accessed 09.03.2022,(in Norwegian).

Ocean Hyway Cluster (2020a). 2030 hydrogen demand in the Norwegian domestic maritime sector. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Ocean Hyway Cluster (2020b). Interactive map - potential maritime hydrogen in Norway. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Owen, S. H. and Daskin, M. S. (1998). Strategic facility location: A review. *European Journal of Operational Research*, 111(3):423–447.

Regjeringen (2021). A european green deal: Norwegian perspectives and contributions. https://www.regjeringen.no/contentassets/38453d5f5f5d42779aaa3059b200a25f/a-european-green-deal-norwegian-perspectives-and-contributions-20.04.2021.pdf. last accessed 09.03.2022.

Rockafellar, R. T. and Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147.

Samferdselsdepartementet (2021). Nasjonal transportplan 2022-2033. https://www.regjeringen.no/no/dokumenter/meld.-st.-20-20202021/id2839503/?ch=1. last accessed 09.03.2022,(in Norwegian).

Santoso, T., Ahmed, S., Goetschalckx, M., and Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115.

Schütz, P., Stougie, L., and Tomasgard, A. (2008). Stochastic facility location with general long-run costs and convex short-run costs. *Computers & Operations Research*, 35(9):2988–3000.

Schütz, P., Tomasgard, A., and Ahmed, S. (2009). Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199(2):409–419.

Shulman, A. (1991). An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations Research*, 39(3):423–436.

Snyder, L. V. (2006). Facility location under uncertainty: a review. *IIE transactions*, 38(7):547–564.

Štádlerová, Š. and Schütz, P. (2021). Designing the hydrogen supply chain for maritime transportation in Norway. In Mes, M., Lalla-Ruiz, E., and Voß, S., editors, *Computational Logistics*, pages 36–50. Springer, Cham.

Statistics Norway (2018). Statistics Norway: 12579: Road traffic volumes. https://www.ssb.no/en/statbank/table/12579/. Online; accessed 02.11.2021.

Ulstein Design (2021). Zero-emission operations in offshore construction market. https://ulstein.com/news/zero-emission-operations-in-offshore-construction-market. last accessed 05.04.2022.

Paper IV

Using Lagrangian relaxation to locate hydrogen production facilities under uncertain demand: A case study from Norway

Šárka Štádlerová, Sanjay Dominik Jena, Peter Schütz

Published in Computational Management Science

Paper 4
Using Lagrangian relaxation to locate hydrogen production facilities under uncertain demand: A case study from Norway

Abstract

Hydrogen is considered a solution to decarbonize the transportation sector, an important step to meet the requirements of the Paris agreement. Even though hydrogen demand is expected to increase over the next years, the exact demand level over time remains a main source of uncertainty. We study the problem of where and when to locate hydrogen production plants to satisfy uncertain future customer demand. We formulate our problem as a two-stage stochastic multi-period facility location and capacity expansion problem. The first-stage decisions are related to the location and initial capacity of the production plants and have to be taken before customer demand is known. They involve selecting a modular capacity with a piecewise linear, convex short-term cost function for the chosen capacity level. In the second stage, decisions regarding capacity expansion and demand allocation are taken. Given the complexity of the formulation, we solve the problem using a Lagrangian decomposition heuristic. Our method is capable of finding solutions of sufficiently high quality within a few hours, even for instances too large for commercial solvers. We apply our model to a case from Norway and design the corresponding hydrogen infrastructure for the transportation sector.

Keywords: Multi-period facility location, Capacity expansion, Uncertain demand, Lagrangian relaxation

5.1 Introduction

According to the emission targets set in the Paris agreement, greenhouse gas emissions (as by 1990) must be decreased by 40% until 2030 (United Nations, 2015). The Norwegian government has set even more ambitious goals regarding the emissions within the transportation sector (Regjeringen, 2019). Specifically, the transition towards zero-emission fuels is a key step in order to meet these targets. The transition from fossil fuels towards hydrogen gained even more importance as countries with diversified energy carrier mix can better handle the current energy crisis in Europe (Crew, 2022). IEA (2022) further states that the global energy crisis accelerates the urgency to use hydrogen, as it contributes to emission reduction targets as well as energy stability. With 92% electricity produced from hydropower, Norway is well positioned to produce green hydrogen, which is required to be produced exclusively from renewable sources using electrolysis (EL).

In Norway, the sector of high-speed passenger ferries and car ferries is operated based on public contracts, and when renewing these contracts, hydrogen can be required as zero-emission fuel (Ocean Hyway Cluster, 2020a). The demand from sectors that are operated based on public contracts may therefore be easier to predict and has a deterministic character as the transition can be forced based on the contracts. There are also alternative zero-emission energy carriers that are relevant in Norway, such as electric batteries and ammonia. However, the future market shares among these fuels are uncertain. Since demand from other relevant sectors such as land-based transport and the offshore sector is highly uncertain, having the ability to expand the production infrastructure is crucial to meet future demand (DNV GL, 2019).

The above motivates our work on the real-world problem of locating hydrogen production facilities in Norway under uncertain demand. The decisions regarding opening location, time and capacity must be taken before the future demand is known. After uncertain demand is disclosed, decisions regarding capacity expansion and production, as well as demand allocation can be taken. The problem formulates as a large mixed-integer programming problem that is, in general, hard to solve. Specifically, as shown in Štádlerová et al. (2022a), the problem can be solved with a commercial solver only for a few scenarios. In this paper, we, therefore, solve this problem using a solution method based on Lagrangian relaxation.

Our contributions are threefold. First, we provide a solution method based on Lagrangian relaxation for the multi-period facility location and capacity expansion problem under uncertainty that allows for solving problems with a sufficiently large number of scenarios within reasonable computing time. Our model formulation includes minimum production requirements motivated by the properties of the production technology for hydrogen. Such requirements can also be found in other industries, for example due to economic or technological considerations (such as minimum batch sizes). Still, our model formulation is general enough to also be applicable if such minimum production requirements do not exist. We compare the performance of the method to the one of Gurobi and discuss the quality of the Lagrangian bound. We further analyze the out-of-sample performance and discuss the value of the stochastic solution. Second, we study the hydrogen production infrastructure for different demand distributions and compare the first-stage solutions to the solution from the expected value problem. The computational results show that the Lagrangian relaxation provides tight lower bounds and that our solution method finds solutions of sufficiently high quality for all tested instances. We further analyze the value of the stochastic solution, indicating that for most problems, the solution of the expected value problem is of no practical use. Third, we analyze the solution obtained for the case of Norway, illustrating the practical usefulness and importance of our approach.

The remainder of this paper is structured as follows: The relevant literature is reviewed in Section 5.2. The mathematical model is introduced in Section 5.3. The solution method is detailed in Section 5.4. The case study is presented in Section 5.5 and the computational results are discussed in Section 5.6. Finally, we conclude in Section 5.7.

5.2 Literature review

We split the literature review in two main parts. In Section 5.2.1, we provide a brief literature review on modelling deterministic and stochastic capacitated facility location problems with piecewise linear costs and/or capacity expansion. We also review facility location and supply chain design problems in the context of hydrogen infrastructure. Solution methods for facility location and supply chain design problems with a focus on two-stage stochastic problems are reviewed in Section 5.2.2.

5.2.1 Capacitated facility location

For an overview on deterministic multi-period facility location and capacity expansion models, we also refer to the reviews by Melo et al. (2009) and Nickel and Saldanha da Gama (2019).

Deterministic multi-period facility location and capacity expansion problems are often modelled with modular capacities. The expansion is then modelled as a jump between available capacity levels and leads to modification of existing facilities (Jena et al., 2015, 2016, 2017; Sauvey et al., 2020; Štádlerová and Schütz, 2021; Štádlerová et al., 2022b). Štádlerová and Schütz (2021) and Štádlerová et al. (2022b) study a problem with modular capacities and piecewise linear short-term production costs (which can be seen as a combination of the problems studied by Correia and Captivo (2003) and Van den Broek et al. (2006)). Similar to Correia and Captivo (2003), they split investment and operational costs and provide specific operational costs to each modular capacity level. However, instead of one unit price for each capacity level, they model a capacity-specific piecewise linear short-term costs function similar to Van den Broek et al. (2006). Van den Broek et al. (2006) combine operational costs depending on installed capacity from Correia and Captivo (2003) with the linear staircase cost approximation from Holmberg (1994). Our modelling approach is identical to Štádlerová and Schütz (2021) and Štádlerová et al. (2022b), as it enables us to model economies and dis-economies of scale in the investment and production processes having modular capacities.

Introducing demand uncertainty is a natural extension of deterministic problems. An early literature review on dynamic facility location and supply chain problems with stochastic parameters can be found in Owen and Daskin (1998). A review on facility location problems under uncertainty is provided by Snyder (2006) and recent summaries on facility location and supply chain problems under uncertainty are presented by Govindan et al. (2017), and Correia and Saldanha da Gama (2019).

Traditionally, two-stage stochastic facility location problems are formulated as single-period problems. An early paper discussing a single-period capacitated facility location problem with random demand and non-linear cost function to model economies of scale is presented by Balachandran and Jain (1976). A generalization of their model is proposed by Schütz et al. (2008) who differentiate between general long-term costs for opening facilities and piecewise linear convex short-term costs for operating facilities. Correia and Melo (2021) study a two-stage multi-period facility location model with capacity expansion and reduction. Due to the complexity of the model, the problem can be solved for only 5 scenarios. The authors further show that using valid inequalities to strengthen the model improves computation times and optimality gaps.

Some supply chain design problems are characterized by a decision structure similar to two-stage facility location problems, as first-stage decisions are related to investments, while the second-stage decisions are related to demand allocation (Lucas et al., 2001). For a review on deterministic, as well as stochastic hydrogen supply chain design problems, we refer to Li et al. (2019). Kim et al. (2008) formulate the model of designing a hydrogen supply chain as a two-stage stochastic mixed-integer problem. Here, the first stage decision is related to investment in production facilities and storage while the second stage decision is related to demand allocation. The work by Almansoori and Shah (2012) and Nunes et al.

(2015) can be seen as an extension of Kim et al. (2008) as the authors consider multiple time periods. Dayhim et al. (2014) present a two-stage stochastic problem for minimizing the total expected daily costs of the hydrogen supply chain facing uncertain demand. Unlike Kim et al. (2008) and Nunes et al. (2015), the authors consider also emission, risk and energy consumption costs. Similar to Nunes et al. (2015), Štádlerová et al. (2022a) present a two-stage multi-period stochastic model to formulate the problem of locating hydrogen facilities. However, the authors extend the model by allowing capacity expansion in the second stage.

5.2.2 Solution methods

Deterministic multi-period facility location and capacity expansion problems are in general hard to solve. The stochastic formulation might be closer to the real-world decision process, but also increases the complexity of the problem, especially when considering integer variables in the second stage. To find quality solutions for large instances, efficient solution algorithms need to be applied.

Lagrangian relaxation combined with heuristics for finding feasible solutions has performed well for deterministic multi-period facility location and capacity expansion problems (see, e.g., Shulman, 1991; Jena et al., 2016, 2017; Štádlerová et al., 2022b). Lagrangian relaxation has also been successfully used to solve single-period stochastic two-stage facility location problems with continuous second-stage variables (see, e.g., Schütz et al., 2008).

Sample average approximation (SAA) improves computational tractability by solving the problem repeatedly with a smaller number of scenarios (Kleywegt et al., 2001). Santoso et al. (2005) combine Benders decomposition with SAA to solve a supply chain design problem with uncertain demand and continuous second-stage variables. Sherali and Zhu (2006) and Angulo et al. (2016) study the application of Benders decomposition for stochastic problems with integer first and second-stage variables.

Nunes et al. (2015) and Štádlerová et al. (2022a) apply SAA to solve the problem of locating hydrogen facilities. Nunes et al. (2015) solve the SAA problems with 15 scenarios. The number of scenarios in Štádlerová et al. (2022a) is limited to 10 as the integer variables in the second stage make the problem harder to solve than the one studied in Nunes et al. (2015). SAA is often used in combination with other solution methods to further improve the quality of the solution (see, e.g., Santoso et al., 2005; Schütz et al., 2009; Li and Zhang, 2018).

Researchers have only recently started to consider uncertainty in multi-period facility location problems with capacity expansion. Correia and Melo (2021) and Štádlerová et al. (2022a) illustrate the challenges when using commercial software to solve two-stage stochastic programming models for this type of problem. To the best of our knowledge, our work is the first to present a solution method based on Lagrangian relaxation for the multi-period facility location problem with uncertain demand and capacity expansion in the second stage.

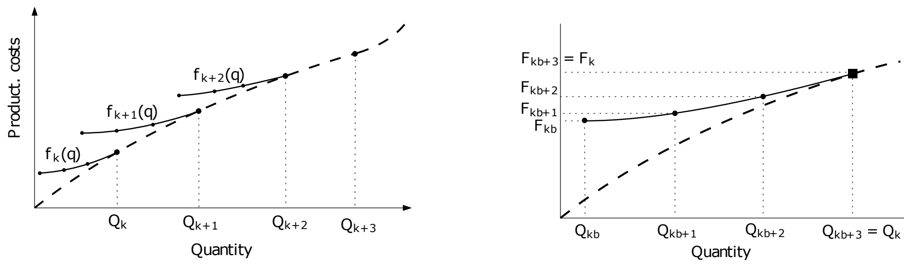
5.3 Mathematical model

We formulate the problem of designing production infrastructure as a two-stage stochastic multi-period facility location and capacity expansion problem with modular capacities. In the first stage, we decide where and when to open new facilities along with their initial capacity levels. Once the demand is known in the second stage, we take decisions related to capacity expansion and demand allocation. The goal is to minimize the expected discounted total costs of satisfying the demand in each scenario.

5.3.1 Problem definition

Candidate locations for production facilities are given by the set $\mathcal{I}$. The investment costs C_{ik} for a new facility depend on location and installed capacity. The feasible production quantity at a facility depends on the installed capacity. The short-term production costs F_{ibkt} then depend on location, installed capacity and its utilization, as well as time period. Customer locations are given by the set $\mathcal{J}$. For each customer, a specific demand, D^s_{jt}, is defined for each time period and each scenario s from the set $\mathcal{S}$. A customer may be served from one or more facilities. However, there are restrictions on which facility can serve which customers. For possible facility-customer combinations, unit distribution costs are based on distance. If the customers' demand D^s_{jt} cannot be satisfied, penalty costs M^D apply for each unit of unsatisfied demand, denoted by d^s_{jt}. Penalty costs for unsatisfied demand can also be interpreted as additional costs for importing the product. If the quantities demanded from a facility are lower than the minimum production quantity for the installed capacity level, penalty costs M^Q for a capacity excess unit q^s_{it} apply. Similar to the penalty cost for unsatisfied demand, the capacity excess costs can be understood as costs for exporting excess production.

Once a facility has been opened, it cannot be closed. However, its capacity may be extended at a later time period to a higher capacity level. The expansion is allowed up to the highest available capacity $\overline{K}$. Capacity expansion leads to facility modification and it represents an expensive strategic decision. Thus, having a relatively short planning horizon, capacity expansion is allowed only once and then, the capacity cannot be changed until the end of the planning horizon $\overline{T}$. Investment costs C_{ik} and expansion costs E_{ikl} represent long-term costs and are separated from short-term production costs. For each capacity level, a specific convex piecewise linear short-term production cost function defines both the cost and the feasible production quantities for the installed capacity. Figure 5.1a exemplifies the link between long-term facility costs and short-term production costs. The short-term production cost function $f_k(q)$ for a specific capacity level k is illustrated in Figure 5.1b, where F_{kb} represents the production costs at a given breakpoint b of the piecewise linear cost function. The lowest breakpoint of the short-term production costs function represents the minimum production requirements for a given capacity level, while the highest breakpoint $\overline{B}_k$ corresponds to the installed capacity and thus to the upper production limit at capacity level k. The upper limit can only be increased by expansion towards a higher capacity level $k + n$. These capacity limits reflect the technological limitations of hydrogen production through electrolysis. This modelling approach is identical to Šťádlerová et al. (2022a), except for the addition of penalties for unsatisfied demand and excess production.



(a) Long-term and short-term production cost functions

(b) Short-term production cost function

Figure 5.1: Long-term and short-term production cost functions

5.3.2 Mathematical formulation

All used sets, parameters and decision variables are summarized below:

Sets

- $\mathcal{B}_k$ Set of breakpoints of the short-term cost function connected to capacity level k, $\mathcal{B}_k = \{1, 2, ..., \overline{B}_k\}$;
- $\mathcal{I}$ Set of candidate locations for production facilities;
- $\mathcal{J}$ Set of customer locations;
- $\mathcal{K}$ Set of available discrete capacity levels, $\mathcal{K} = \{1, 2, ..., \overline{K}\}$;
- $\mathcal{S}$ Set of scenarios;
- $\mathcal{T}$ Set of time periods, $\mathcal{T} = \{1, 2, ..., \overline{T}\}$;
- $\mathcal{T}_1$ Set of time periods corresponding to the first stage, $\mathcal{T}_1 \subset \mathcal{T}$.

Parameters and coefficients

C_{ik}	Investment costs at location $i \in \mathcal{I}$ for capacity level $k \in \mathcal{K}$;
D_{jt}^s	Demand at customer location $j \in \mathcal{J}$ in period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$;
E_{ikl}	Costs of expanding at facility $i \in \mathcal{I}$ from capacity level $k \in \mathcal{K}$ to capacity level $l \in \mathcal{K} : l > k$;
F_{ibkt}	Production costs at facility $i \in \mathcal{I}$ at breakpoint $b \in \mathcal{B}_k$ at the short-term cost function of capacity level $k \in \mathcal{K}$ in period $t \in \mathcal{T}$;
L_{ij}	1 if demand at location $j \in \mathcal{J}$ can be served from facility $i \in \mathcal{I}$, 0 otherwise;
Q_{bk}	Production volume at breakpoint $b \in \mathcal{B}_k$ of the short-term cost function, for capacity level $k \in \mathcal{K}$;
T_{ij}	Distribution costs from facility $i \in \mathcal{I}$ to customer $j \in \mathcal{J}$;
M^D	Penalty costs for one unit of unsatisfied demand;
M^Q	Penalty costs for one excess unit;
y_{ikk0}	1 if a facility is opened at location $i \in \mathcal{I}$ with capacity level $k \in \mathcal{K}$ at the beginning of the planning horizon, 0 otherwise;
δ_t	Discount factor in period $t \in \mathcal{T}$;
p^s	Probability of scenario $s \in \mathcal{S}$.

Decision variables

The mathematical model uses the following decision variables:

d_{jt}^s	Shortfall variable: amount of not satisfied demand at customer location $j \in \mathcal{J}$ in period $t \in \mathcal{T}$ for scenario $s \in \mathcal{S}$;
q_{it}^s	Capacity excess variable: amount of production excess units from facility location $i \in \mathcal{I}$ in period $t \in \mathcal{T}$ for scenario $s \in \mathcal{S}$ that is not distributed to customers;
x_{ijkt}^s	Amount of customer demand at location $j \in \mathcal{J}$ satisfied from facility $i \in \mathcal{I}$ operating at capacity level $k \in \mathcal{K}$ in period $t \in \mathcal{T}$ in scenarios $s \in \mathcal{S}$;
y_{iklt}^s	1 if facility is operated at location $i \in \mathcal{I}$, originally opened at capacity level $k \in \mathcal{K}$, and operating at capacity level $l \in \mathcal{K} : l \geq k$ in period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$, 0 otherwise;
μ_{bilt}^s	Weight of breakpoint $b \in \mathcal{B}_l$ at location $i \in \mathcal{I}$ for capacity level $l \in \mathcal{K}$ in period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$.

We present a two-stage stochastic multi-period Mixed-Integer-Programming (MIP) model. The model is similar to Štádlerová et al. (2022a), but additionally provides relatively complete recourse, as we introduce variables for unsatisfied demand and capacity excess. The model is given as:

$$\min \sum_{s\in\mathcal{S}} p^s \Bigg[\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \delta_t C_{ik} \left(y^s_{ikkt} - y^s_{ikk(t-1)} \right) + \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}:l>k} \sum_{t\in\mathcal{T}} \delta_t E_{ikl} \left(y^s_{iklt} - y^s_{ikl(t-1)} \right) + \sum_{b\in\mathcal{B}} \sum_{i\in\mathcal{I}} \sum_{l\in\mathcal{K}} \sum_{t\in\mathcal{T}} \delta_t F_{iblt} \mu^s_{bilt} + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{l\in\mathcal{K}} \sum_{t\in\mathcal{T}} \delta_t T_{ij} x^s_{ijlt} \sum_{j\in\mathcal{J}} \sum_{t\in\mathcal{T}} M^D d^s_{jt} + \sum_{i\in\mathcal{I}} \sum_{t\in\mathcal{T}} M^Q q^s_{it} \Bigg] \tag{5.1}$$

subject to:

$$\sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}:l\geq k} y^s_{iklt} \leq 1, \qquad i\in\mathcal{I}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{5.2}$$

$$\sum_{k\in\mathcal{K}} \sum_{l\in\mathcal{K}:l>k} y^s_{iklt} = 0, \qquad i\in\mathcal{I}, t\in\mathcal{T}_1, s\in\mathcal{S}, \tag{5.3}$$

$$\sum_{t'=1}^{t-1} y^s_{ikkt'} \geq \sum_{l\in\mathcal{K}:l>k} y^s_{iklt}, \qquad i\in\mathcal{I}, k\in\mathcal{K}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{5.4}$$

$$\sum_{l\in\mathcal{K}:l\geq k} y^s_{iklt} \geq \sum_{l\in\mathcal{K}:l\geq k} y^s_{ikl(t-1)}, \qquad i\in\mathcal{I}, k\in\mathcal{K}, t\in\mathcal{T}, s\in\mathcal{S} \tag{5.5}$$

$$y^s_{iklt} - y^s_{ikl(t-1)} \geq 0, \qquad i\in\mathcal{I}, k\in\mathcal{K}, l\in\mathcal{K}: l>k, t\in\mathcal{T}, s\in\mathcal{S}, \tag{5.6}$$

$$\sum_{b\in\mathcal{B}_l} \mu^s_{bilt} = \sum_{k\in\mathcal{K}} y^s_{iklt}, \qquad i\in\mathcal{I}, l\in\mathcal{K}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{5.7}$$

$$\sum_{j\in\mathcal{J}} \sum_{l\in\mathcal{K}} x^s_{ijlt} + q^s_{it} = \sum_{b\in\mathcal{B}_l} \sum_{l\in\mathcal{K}} Q_{bl} \mu^s_{bilt}, \qquad i\in\mathcal{I}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{5.8}$$

$$\sum_{i \in \mathcal{I}} \sum_{l \in \mathcal{K}} x^s_{ijlt} + d^s_{jt} = D^s_{jt}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}, \quad (5.9)$$

$$x^s_{ijlt} \leq L_{ij} D^s_{jt} \sum_{k \in \mathcal{K}: k \leq l} y^s_{iklt}, \qquad i \in \mathcal{I}, j \in \mathcal{J}, l \in \mathcal{K}, t \in \mathcal{T}, s \in \mathcal{S}, \quad (5.10)$$

$$\frac{1}{|\mathcal{S}|} \sum_{s' \in \mathcal{S}} \sum_{l \in \mathcal{K}: l \geq k} y^{s'}_{iklt} = \sum_{l \in \mathcal{K}: l \geq k} y^s_{iklt}, \qquad i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, s \in \mathcal{S}, \quad (5.11)$$

$$y^s_{iklt} \in \{0, 1\}, \qquad i \in \mathcal{I}, k \in \mathcal{K}, l \in \mathcal{K} : l \geq k, t \in \mathcal{T}, s \in \mathcal{S}, \quad (5.12)$$

$$x^s_{ijlt} \geq 0, \qquad i \in \mathcal{I}, j \in \mathcal{J}, l \in \mathcal{K}, t \in \mathcal{T}, s \in \mathcal{S} \quad (5.13)$$

$$\mu^s_{bilt} \geq 0, \qquad b \in \mathcal{B}_l, i \in \mathcal{I}, k \in \mathcal{K}, t \in \mathcal{T}, s \in \mathcal{S}, \quad (5.14)$$

$$q^s_{it} \geq 0, \qquad i \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}, \quad (5.15)$$

$$d^s_{jt} \geq 0, \qquad j \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}. \quad (5.16)$$

Objective (5.1) minimizes the expected discounted sum of investment, expansion, production, and distribution costs as well as the penalty costs for unsatisfied demand and excess capacity. Constraints (5.2) state that for each time period and scenario, at most one facility can be operated at a given location. Constraints (5.3) ensure that in the first stage, facilities can be only opened, but not expanded, while Inequalities (5.4) only allow expansion of opened facilities. Constraints (5.5) ensure that once a facility is opened, it cannot be closed, but only expanded, while Constraints (5.6) require that an open facility can only be expanded once during the planning horizon. Equalities (5.7) link capacity level k with the appropriate short-term cost function and ensure that only opened facilities can be used for production. Constraints (5.8) ensure that the whole production is either distributed to customers or allocated to the capacity excess variable. The constraints also implicitly assure the minimum production requirements through the quantity Q_{bl} given by the smallest breakpoint b. Note that this formulation is also applicable for problems without minimum production requirements, as we can define the quantity belonging to the smallest breakpoint as zero. Equations (5.9) ensure that demand is satisfied or registered as demand shortfall. Restrictions (5.10) are formulated in the form of strong inequalities. They limit which facility can satisfy which customer and link the distribution variable to the operated capacity level. Such linking constraints provide stronger bounds and lead to lower integrality gaps from linear relaxation than aggregated linking constraints (see, e.g., Jena et al., 2016). Constraints (5.11) are the non-anticipativity constraints that ensure that the opening capacity level k is the same for all scenarios while the operating capacity level l is scenario specific. Constraints (5.12)–(5.16) are the non-negativity and binary requirements.

5.4 Lagrangian relaxation

In the domain of facility location, Lagrangian relaxation has mostly been applied in deterministic settings (see, e.g., Shulman, 1991; Jena et al., 2016, 2017; Šťádlerová et al., 2022b). Given the similar structure of the here considered facility location problem, Lagrangian relaxation remains an attractive candidate, even when considering multiple demand scenarios. We now present the Lagrangian heuristic used to solve our stochastic problem. Specifically, we relax demand constraints (5.9) which are the only constraints connecting the decision variables among the different facility locations and have been a popular choice in the literature (Shulman, 1991; Schütz et al., 2008; Jena et al., 2016). We define λ^s_{jt} as the matrix of Lagrangian multipliers and we obtain the following Lagrangian subproblem:

$$\begin{aligned} LR(\boldsymbol{\lambda}) = \min \sum_{s\in\mathcal{S}} p^s \Bigg[& \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} \delta_t C_{ik}\left(y^s_{ikkt} - y^s_{ikk(t-1)}\right) + \\ & \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}:l>k}\sum_{t\in\mathcal{T}} \delta_t E_{ikl}(y^s_{iklt} - y^s_{ikl(t-1)}) + \\ & \sum_{i\in\mathcal{I}}\sum_{b\in\mathcal{B}_l}\sum_{l\in\mathcal{K}}\sum_{t\in\mathcal{T}} \delta_t F_{iblt}\mu^s_{bilt} + \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} M^Q q^s_{it} + \\ & \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{K}}\sum_{t\in\mathcal{T}} (\delta_t T_{ijt} - \lambda^s_{jt}) x^s_{ijlt} + \\ & \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}} (M^D - \lambda^s_{jt}) d^s_{jt} + \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}} \lambda^s_{jt} D^s_{jt} \Bigg], \end{aligned} \tag{5.17}$$

subject to Constraints (5.2)–(5.8) and (5.10)–(5.15). In the relaxed problem, the variable d_{jt} is unbounded and it has no connection to any other decision variable. Since we have a minimization problem, it can be shown that the term $\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}(M^D-\lambda^s_{jt})d^s_{jt}$ becomes zero in any optimal solution and can hence be omitted. Further, for given multipliers λ^s_{jt}, the expression $\sum_{s\in\mathcal{S}}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}p^s\lambda_{jt}D^s_{jt}$ is constant. As all constraints are defined separately for each facility location $i\in\mathcal{I}$, we can decompose the problem and solve it independently for each facility location $i\in\mathcal{I}$. We can then define $LR(\lambda)=\sum_{i\in\mathcal{I}}g_i(\lambda)+\sum_{s\in\mathcal{S}}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}p^s\lambda_{jt}D^s_{jt}$ where $g_i(\lambda)$ is the optimal value of the Lagrangian subproblem for location i:

$$
\begin{aligned}
g_i(\lambda) = \min \sum_{s\in\mathcal{S}} p^s \Bigg[& \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} \delta_t C_{ik}\left(y^s_{ikkt}-y^s_{ikk(t-1)}\right)+ \\
& \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{K}:l>k}\sum_{t\in\mathcal{T}} \delta_t E_{ikl}(y^s_{iklt}-y^s_{ikl(t-1)})+ \\
& \sum_{b\in\mathcal{B}_l}\sum_{l\in\mathcal{K}}\sum_{t\in\mathcal{T}} \delta_t F_{iblt}\mu^s_{bilt}+\sum_{t\in\mathcal{T}} M^Q q^s_{it}+ \\
& \sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{K}}\sum_{t\in\mathcal{T}}(\delta_t T_{ijt}-\lambda_{jt})x^s_{ijlt}\Bigg],
\end{aligned}
\tag{5.18}
$$

subject to constraints (5.2)-(5.8) and (5.10)-(5.15) defined for the specific facility $i\in\mathcal{I}$.

5.4.1 Solving the Lagrangian subproblem

The optimal solution to the Lagrangian subproblem represents the optimal opening and expansion schedule and capacity level for each facility all scenarios and over all scenarios such that the expected total costs (5.18) are minimized. In deterministic settings, such schedules have been found by solving a shortest path problem via dynamic programming (see, e.g., Shulman, 1991; Jena et al., 2016; Štádlerová et al., 2022b). Given that, in our two-stage stochastic problem, the expansion schedule (i.e, the second-stage decisions) may be different for each scenario, shortest path networks including all opening and expansion decisions would be too complex and computationally intractable. Our approach, therefore, evaluates the optimal expansion schedule for all possible opening decisions separately. Specifically, for each opening capacity level and time period (i.e, the first-stage decisions), the shortest path problem is solved via dynamic programming independently for each scenario starting from the defined opening time period and capacity level, similar to Shulman (1991), Jena et al. (2016) and Štádlerová et al.

(2022b). For each scenario, at most one capacity expansion is allowed. The shortest path problem for solving the Lagrangian subproblem is detailed in Section 5.4.1. For a given capacity level, time period, and scenario, the problem of customer allocation then becomes a continuous knapsack problem which is explained next.

Continuous knapsack

The costs of the optimal demand allocation for a given capacity level $l \in \mathcal{K}$, time period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$ can be computed by solving a continuous knapsack problem with piecewise linear costs (Amiri, 1997; Christensen and Klose, 2021). The costs of the continuous knapsack for a given capacity level l consist of production costs, penalty costs for capacity excess and reduced distribution costs. The costs are calculated for a given facility $i \in \mathcal{I}$, capacity level l, time period t and scenario s. Since the continuous knapsack is calculated for a given capacity level l, the strong inequalities (5.10) are considered in the calculation of the knapsack costs. We formulate the continuous knapsack problem as:

$$K_{ilt}^{s}(\boldsymbol{\lambda}) = \min \sum_{b \in \mathcal{B}_l} F_{iblt}\mu_{bilt}^{s} + M_q q_{it}^{s} + \sum_{j \in \mathcal{J}} (T_{ij} - \lambda_{jt}^{s}) x_{ijlt}^{s}, \tag{5.19}$$

subject to:

$$x_{ijlt}^{s} \leq L_{ij} D_{jt}^{s}, \qquad j \in \mathcal{J}, \tag{5.20}$$

$$\sum_{j \in \mathcal{J}} x_{ijt}^{s} + q_{it}^{s} = \sum_{b \in \mathcal{B}_l} Q_{bl}\mu_{bilt}^{s}, \tag{5.21}$$

$$\sum_{b \in \mathcal{B}_l} \mu_{bilt}^{s} = 1, \tag{5.22}$$

$$q_{it}^{s} \geq 0, \tag{5.23}$$

$$x_{ijlt}^{s} \geq 0, \qquad j \in \mathcal{J}, \tag{5.24}$$

$$\mu_{bilt}^{s} \geq 0, \qquad b \in \mathcal{B}_l. \tag{5.25}$$

The problem (5.19) – (5.25) is similar to the one solved by Schütz et al. (2008). However, in contrast to Schütz et al. (2008), we have a minimum production requirement for each capacity level and allow for capacity excess. For a given capacity level $l \in \mathcal{K}$, period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$, we calculate the unit production costs as $u_{ilbt} = \frac{F_{ib+1lt} - F_{iblt}}{Q_{b+1l} - Q_{bl}}$. We further define the marginal costs of serving one additional demand unit as: $m^s_{ijlbt} = T_{ij} - \lambda^s_{ij} + u_{ilbt}$. Note that the marginal costs are dependent on the line-piece of the short-term cost function. For each customer, we calculate the reduced costs $T_{ij} - \lambda^s_{ij}$ and start allocating customers with the lowest reduced costs until $m^s_{ijlbt} > 0$ for the first time or until the capacity limit of the line-piece Q_{b+1l} is reached. For the next line-piece, the marginal costs must be updated. However, the ordering of customers according to their reduced costs remains unchanged. We continue adding customers until $m^s_{ijlbt} > 0$ or until the capacity limit $Q_{\overline{B}_l}$ is reached.

If the minimum production requirement for a given capacity cannot be fulfilled with customers with negative reduced costs, we may also have to add customers with positive reduced costs. Assuming that penalty costs are always higher than the costs of satisfying customers with positive reduced costs, we prefer customers with positive reduced costs to using variables q^s_{it}. However, if there are no more customers that could be added and the minimum production requirement is still not satisfied, we can use variables q^s_{it} that allow satisfying the minimum production requirement for penalty costs. If the penalty costs are sufficiently high, a capacity decision leading to $q^s_{it} > 0$ will most likely not be optimal since demand does not need to be satisfied in the relaxed problem.

Formulating the shortest path problem

As previously mentioned, in deterministic problems, the problem of finding the optimal opening and expansion decision can be formulated as a shortest path problem in a single graph and solved via dynamic programming (Shulman, 1991; Jena et al., 2016; Štádlerová et al., 2022b). In our scenario-based stochastic problem, such a single graph formulation is not suitable, since the opening decision has to remain the same for all scenarios, but the expansion decision can be different for each scenario. Therefore, we define one shortest path problem for each tuple $(\overline{k}_0, \overline{t}_0)$ of opening capacity level $\overline{k}_0 \in \mathcal{K} \cup \{0\}$ and opening time period $\overline{t}_0 \in \mathcal{T}$. For each given $(\overline{k}_0, \overline{t}_0)$, the second stage problem is then separable in scenarios and we can calculate the shortest path problem separately for each scenario. Finally, we choose the first-stage opening decision that leads to the lowest expected costs over the shortest path problems.

For given opening decision $(\overline{k}_0, \overline{t}_0)$, let $C^E(\overline{k}_0, \overline{t}_0)$ denote the costs of the expected shortest path. The costs of opening and operating a facility during the opening period are equal to the investment costs and the expected costs of the continuous knapsack: $\delta_{t_0} C_{\overline{k}_0 \overline{t}_0} + \sum_{s \in \mathcal{S}} \delta_{t_0} p^s K^s_{i\overline{k}_0\overline{t}_0}(\boldsymbol{\lambda})$. The total costs can then be written as: $\delta_{t_0} C_{i\overline{k}_0 \overline{t}_0} + \sum_{s \in \mathcal{S}} \delta_{t_0} p^s K^s_{i\overline{k}_0\overline{t}_0}(\boldsymbol{\lambda}) + C^E(\overline{k}_0\overline{t}_0)$.

For given opening decision $(\overline{k}_0, \overline{t}_0)$ and given scenario $s \in \mathcal{S}$, the graph structure is illustrated in Figure 5.2. Let $l_{\overline{T}}$ denote the capacity level at the end of the planning horizon. The graph shows that after investing in capacity level $\overline{k}_0$, we can either keep the capacity at level $\overline{k}_0$ or we can expand once during the planning horizon towards a higher capacity level $l_{\overline{T}} \in \mathcal{K} : l_{\overline{T}} > \overline{k}_0$. Note that all capacities larger than $\overline{k}_0$ are available for expansion. However, we are not allowed to reduce the capacity level below the level given by $\overline{k}_0$.

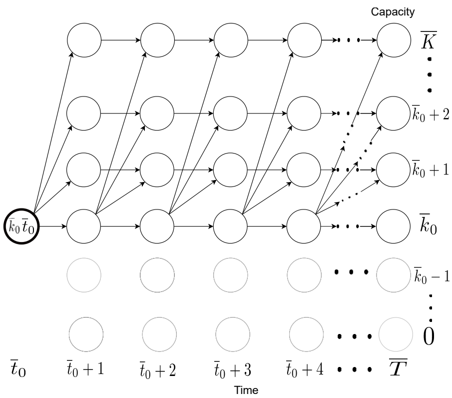


Figure 5.2: Structure of the shortest path problem for a given investment decision and scenario

The costs for an arc from node $(k, t-1)$ to node (k', t) in our graph are given as:

$$C(k,t-1)(k',t) = \begin{cases} E_{ikk'} + K_{ik't}(\lambda) & \text{if} \quad k = \overline{k}_0 \wedge k' = l_{\overline{T}}, & (5.26) \\ K_{ik't}(\lambda) & \text{if} \quad k = k', & (5.27) \\ +\infty & \text{else.} & (5.28) \end{cases}$$

Equation (5.26) calculates the costs of expanding a facility as the sum of expansion costs $E_{ikk'}$ and the costs of continuous knapsack $K_{ik't}(\boldsymbol{\lambda})$. Equation (5.27) calculates the costs of operating the facility if there is no change in the installed capacity level. The short-term production costs are then given as the costs of the continuous knapsack $K_{ik't}(\boldsymbol{\lambda})$. We define the costs of all other combinations as $+\infty$ (5.28) as these are infeasible and hence can be omitted in the graph structure.

5.4.2 Updating the Lagrangian multipliers

The lower bound on the Objective (5.1) is given by solving (5.17) subject to Constraints (5.2)–(5.8) and (5.10)–(5.15) for given multipliers λ^s_{jt}. In order to find the highest possible lower bound, we have to find a $\boldsymbol{\lambda}$ that maximizes the Lagrangian dual problem: $LD = \max_\lambda LR(\boldsymbol{\lambda})$. To solve the LD problem, we iteratively use the box step method (Marsten et al., 1975) similar to Schütz et al. (2009) and Šťádlerová et al. (2022b), as this method allows us to update the multipliers without computing an upper bound. We calculate the subgradient ∇^{ms}_{jt} as $\nabla^{ms}_{jt} = D^s_{jt} - \sum_{i\in\mathcal{I}} x^s_{ijt}$ in each iteration m and for each scenario s. We then define $L^m = LR(\boldsymbol{\lambda}^m) - \sum_{j\in\mathcal{J}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} p^s \lambda^{ms}_{jt} \nabla^{ms}_{jt}$ and find the updated multipliers by solving the following linear optimization problem:

$$\max \phi \tag{5.29}$$

$$\phi \leq L^i + \sum_{j\in\mathcal{J}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{S}} p^s \nabla^{is}_{jt} \lambda^{m+1,s}_{jt}, \qquad i = 1, ..., m, \tag{5.30}$$

$$\lambda^{m+1,s}_{jt} \leq \lambda^{ms}_{jt} + \Delta^{ms}_{jt}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}, \tag{5.31}$$

$$\lambda^{m+1,s}_{jt} \geq \lambda^{ms}_{jt} - \Delta^{ms}_{jt}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}, \tag{5.32}$$

$$\phi \in \mathbb{R}, \lambda^{m+1,s}_{jt} \in \mathbb{R}. \tag{5.33}$$

We limit how much the Lagrangian multipliers can change in each iteration using box constraints (5.31) and (5.32). These boxes are specific for each variable λ^s_{jt}. If the sign of the subgradient ∇^{ms}_{jt} changes from the previous iteration $m-1$, we decrease the box size as: $\Delta^{ms}_{jt} = \alpha \Delta^{ms}_{jt}$, where $0 < \alpha < 1$ (Šťádlerová et al., 2022b). The aim of reducing the box size is to speed up the procedure of finding the optimal multipliers (Marsten et al., 1975). If the multipliers do not change for three consecutive iterations, we reset the box size and allow large changes of the multipliers again in order to escape a local optimum.

5.4.3 Upper bound

We use a greedy heuristic to build a feasible solution based on the solution of the relaxed problem (i.e., the LD). Due to capacity excess and shortfall variables, the solution to the relaxed problem is always feasible. However, these variables may imply high penalty costs. In our upper bound heuristic, we aim to find first-stage solutions that are feasible in all scenarios without or with minimal penalty costs. The heuristic is an extension of the deterministic solution method presented by Šťádlerová et al. (2022b). The main steps of the heuristic are illustrated in Figure 5.3.

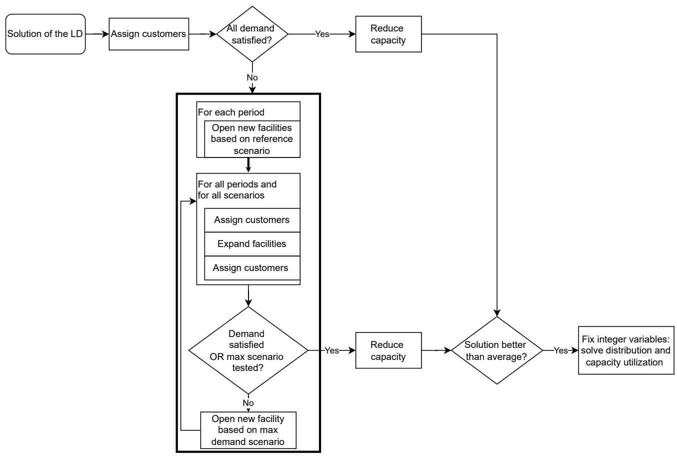


Figure 5.3: Upper bound structure

We initialize the solution using the installed capacity from the *Solution of the LD*, i.e., the capacity level of opened facilities. The allocation and distribution decisions from the relaxed problem are ignored when we assign customers to facilities. Note that the solution of the relaxed problem satisfies the non-anticipativity constraints, so customers can be assigned to facilities separately for each scenario.

In step *Assign customers*, for a given scenario and time period, we create pairs of available facilities $i \in \mathcal{I}$ and unsatisfied customers $j \in \mathcal{J}$. These pairs are sorted in increasing order of their reduced transportation costs $T_{ij} - \lambda^s_{jt}$. We start with the pair with the lowest reduced transportation costs and serve the unsatisfied customer from the corresponding facility. We repeat this step until all available capacity is used or the demand of all customers is satisfied.

Within step *Assign customers*, we also verify the minimum production requirements and try to fix them. If the minimum production requirements of facility B are not satisfied, the heuristic selects a facility with high utilization A. If there are customers that can be satisfied both from A and B, the heuristic uses facility A to shift some of its production to facility B until the production is sufficient. Otherwise, we have to find facility C, which has common customers with both

A and B. Then, we shift some production from facility A via auxiliary facility C to facility B. The heuristic uses up to three auxiliary facilities to shift production between A and B. When shifting the production quantities from A to B, customers are sorted in increasing order based on their reduced transportation costs from facility B. We start reallocating customers with the lowest reduced transportation costs. After reallocation, the spare capacity in facility A is used to satisfy additional unsatisfied customers.

The capacity obtained from the solution of the relaxed problem is most likely not sufficient to satisfy demand in all scenarios. If there are unsatisfied customers after step *Assign customers* considering *Solution of the LD*, the upper bound heuristic increases the capacity to satisfy all customers or to minimize the penalty costs for demand shortfall. These steps are illustrated in Figure 5.3 in the bold frame.

In general, a new facility can be opened at a location without a facility. Expansion is allowed only at a location with an existing facility that has not been expanded yet. When selecting the facility that has to be opened or expanded, there are usually several candidates. We choose the candidate that can satisfy most of the unsatisfied customers. In case of a tie, we prioritize the facility with lower production costs. The chosen capacity for opening or expanding a facility is the lowest possible capacity level that can satisfy the demand.

We execute the upper bound heuristic repeatedly for 4 different reference scenarios: maximum, minimum, mean and median demand scenarios and then, we select the solution with the lowest objective. Since the opening decision has to be equal for all scenarios, we start with step *Open new facilities based on reference scenario* and implement the first-stage decisions based on the chosen reference scenario for all other scenarios before executing routines that are specific for each scenario. Considering only 4 reference scenarios enables shorter computation times compared to evaluating first-stage decisions of each scenario. Simultaneously, first-stage decisions provided by one of the reference scenarios have shown to be sufficiently good for our upper bound.

After the opening decisions are fixed, step *Assign customers* can be again performed for each scenario independently as well as the expansion decisions in the step *Expand facilities* since these are the second-stage decisions. If the installed capacity is still not sufficient and our reference scenario differs from the maximum demand scenario, the heuristics performs the step *Open new facility based on max demand scenario*, where the opening decisions are taken based on unsatisfied customers in the scenario with maximum demand. Then, the capacity installed in the first stage increases in all scenarios. Note that these new facilities can later be expanded as well.

The upper bound heuristic aims to install sufficient capacity to avoid penalties for demand shortfall. However, the solution of the LD as well as in the upper bound heuristic may have installed more capacity than necessary. Therefore, we try to reduce the installed capacity or remove some facilities in order to improve the total costs. We first try to remove facilities with capacity excess. Specifically, we identify a facility that causes penalties for capacity excess and check whether the allocated customers can be served from other opened facilities in all time periods and scenarios. These facilities need to have some spare capacity and satisfy the distance limit to the customers. If all customers can be reallocated, we remove the facility. Further, we extend the deterministic procedure from Jena et al. (2016) to our stochastic problem. We fix the demand allocation decisions, and use a dynamic programming algorithm, to find optimal opening and expansion capacities and time periods to satisfy the given quantities.

If the obtained solution is better than the average of the previously found solutions, we fix all integer variables and solve a problem consisting of demand allocation and facility utilization with Gurobi. When evaluating the average costs, we consider objectives before re-optimizing distribution and facility utilization. This enables us to save time, as we do not need to re-optimize all available solutions and reduce the risk of ignoring a potentially good solution.

5.5 Case study

In this section, we introduce the input data for our case study from Norway. The case reflects the real-world problem of producing hydrogen for the Norwegian transport sector. However, our model is applicable to variety of facility location and capacity expansion planning problems.

5.5.1 Candidate locations and production costs

We consider 17 ports along the Norwegian coast as candidate locations derived from the interactive map provided by Ocean Hyway Cluster (2020b). For testing purposes, we further extend the number of candidate locations to 34. All these locations are Norwegian ports and contain the original 17 locations as a subset.

In our case study, we consider alkaline electrolysis as production technology since it is as of today the most mature and the cheapest available technology. For alkaline electrolysers, minimum production requirements must be considered (Andrenacci et al., 2022). We assume that investment costs are the same for all facility locations. We approximate the long-term production cost function with 8 and 16 modular capacity levels, each with specific investment costs. For 16 capacity levels, each of the original 8 capacity levels is split into two levels. The investment costs for 8 capacity levels are shown in Table 5.1. All investment costs are calculated based on the model by Jakobsen and Åtland (2016).

Discrete capacity	1	2	3	4	5	6	7	8
Capacity [tonnes/day]	0.6	3.1	6.2	12.2	30.3	61.0	151.5	304.9
Investment [mill. €]	1.4	6.0	11.2	20.5	46.5	87.2	197.7	371.5

Table 5.1: Investment costs for electrolysis (Štádlerová and Schütz, 2021)

Electrolysis production costs are highly dependent on electricity prices. Even though Norway is split into 5 electricity price regions, the prices differ mainly between the northern part and the southern part of Norway. We, therefore, use different production costs dependent on whether the candidate locations are in northern Norway (N) or southern Norway (S). All candidate locations situated in and north of Trondheim are considered to belong to the northern region.

To calculate the production costs in periods 1 to 9, we use the 2021 yearly average electricity prices for the two regions. On average, the prices in the southern region were 1.8 times higher than prices in the northern region in 2021 (Nord Pool AS, 2022). For periods 10 to 14, we use the electricity price based on the forecast from NVE (2021) that predicts a smaller difference between the northern and the southern region. According to this forecast, the price in the southern region should be about 1.2 times higher than in the northern region. The production costs are calculated using the model by Jakobsen and Åtland (2016). The production costs at 100% capacity utilization for southern (S) and northern Norway (N) are shown in Table 5.2.

Discrete capacity	1	2	3	4	5	6	7	8
Capacity [tonnes/day]	0.6	3.1	6.2	12.2	30.3	61.0	151.5	304.9
Production S[€/kg]	4.26	4.21	4.20	4.18	4.16	4.14	4.13	4.11
Production N[€/kg]	2.54	2.50	2.47	2.46	2.44	2.42	2.40	2.39

Table 5.2: Production costs for EL at 100% capacity utilization

For each capacity level, we approximate the short-term production cost function by a piecewise linear function with 4 breakpoints. The production range for electrolysis is 15% – 100% (NEL Hydrogen, 2018). Thus, we define breakpoints of the short-term cost function at 15%, 50%, 80%, and 100% of the installed capacity level. For each capacity level, the 15% breakpoint represents the minimum production requirement.

5.5.2 Penalty costs

We define penalty costs for each unit of demand shortfall and capacity excess. Since the focus of this case study is on domestic hydrogen production for domestic customers, we set high penalties for both demand shortfall and capacity excess of 10^9 €/kg to avoid both import and export of hydrogen.

5.5.3 Distribution costs

Hydrogen is distributed in trucks as pipelines are not a suitable distribution solution for Norway. Distribution costs per kilometer and kilogram of hydrogen are based on the hydrogen distribution study provided by Danebergs and Aarskog (2020) and taken from Štádlerová and Schütz (2021). The distribution costs are defined for different distance intervals as shown in Table 5.3. The maximum distance between production facilities and customers is 1000 km.

Distance [km]	1-50	51-100	101-200	201-400	401-800	801-1000
Costs	0.00498	0.00426	0.00390	0.00372	0.00363	0.00360

Table 5.3: Hydrogen distribution costs in [€/km/kg H_2] (Štádlerová and Schütz, 2021)

5.5.4 Demand

The total hydrogen demand consists of three components:

- Maritime demand (Ocean Hyway Cluster, 2020a),
- Land-based demand (DNV GL, 2019),
- Offshore demand (Aglen and Hofstad, 2022).

The maritime demand is based on public contracts for high-speed passenger ferries and car ferries. This component is considered to be deterministic and is present in all demand scenarios. In the land-based and offshore sectors, the future demand share among competing zero-emission carriers is highly uncertain. Thus, the demand share, and as such the total demand for hydrogen, from these sectors differs in each scenario. The deterministic demand estimations in DNV GL (2019) and Aglen and Hofstad (2022) are highly uncertain and we consider them to represent the maximal potential demand for hydrogen in these sectors.

Figure 5.4 shows the evolution of the maximum potential demand for all demand components over the planning horizon. Maritime demand is characterized by a steady demand increase, and the demand jump in time period 11 represents ships on the coastal route Bergen-Kirkenes that will switch to hydrogen fuel as well. The transition towards hydrogen in the land-based sector is expected to come in two waves that cause the demand jumps in periods 4 and 9. In the offshore sector, most of the ships should be transformed to use hydrogen fuel before period 10. Afterwards, the demand is almost constant.

We assume that the share of each of the demand components is independent and given by a specific distribution. The effect of competing zero-emission is

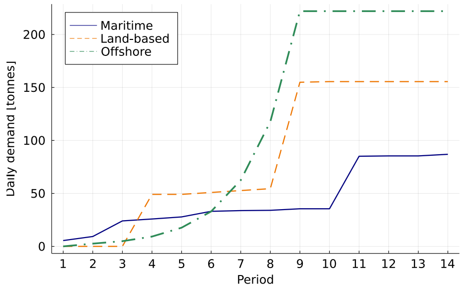


Figure 5.4: Annual daily demand

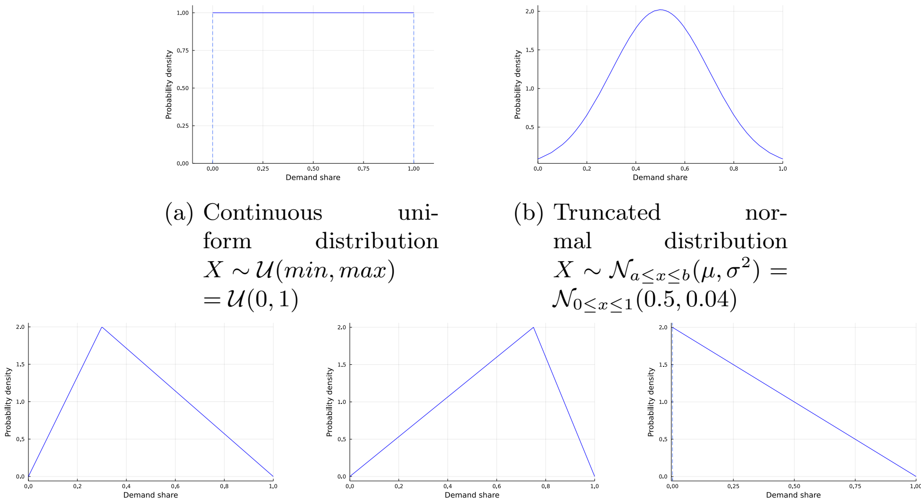
considered by using scenarios with low hydrogen demand since some customer segments can decide for using battery electric solutions or ammonia instead of hydrogen. Since also the demand distribution is subject to uncertainty, we study the impact of different demand distributions on the infrastructure.

To assess the impact the demand distribution has on the solution, we solve our model for a uniform (unif) distribution, a normal (norm) distribution, as well as three different triangular distributions, see Figure 5.5. The uniform and the normal distributions have identical expected values, whereas the expected value of the triangular distributions depends on their shape. The left-skewed triangular distribution (trg-L) assumes demand to consist mainly of maritime demand and a low share of land-based and offshore demand, while the right-skewed triangular distribution (trg-R) assumes a high overall demand level. Finally, we present a left-skewed triangular distribution with the expected value equal to the maritime demand level with a very low of the Land-based and Offshore demand (trg-min).

We consider aggregated daily demand in 70 and 390 demand points located in Norway. For the maritime and offshore sectors, there are 51 demand points located in Norwegian ports. For the instances with 70 customers, we consider additional 19 municipalities with the highest road traffic volumes (Statistics Norway, 2018). Road traffic demand is then divided among the 70 customers according to the relative traffic volume. For the instances with 390 demand points, we divide road traffic demand among the 390 municipalities. Note that municipalities with a daily hydrogen demand from road traffic of less than 10 kg are neglected.

5.6 Computational results

All calculations have been carried out on a Linux cluster with two 3.6 GHz Intel Xeon Gold 6244 CPU (core) processors and 384 GB RAM. We use commercial software Gurobi Optimizer 9.5. to solve the demand allocation and facility uti-



(a) Continuous uniform distribution $X \sim \mathcal{U}(min, max) = \mathcal{U}(0, 1)$

(b) Truncated normal distribution $X \sim \mathcal{N}_{a \leq x \leq b}(\mu, \sigma^2) = \mathcal{N}_{0 \leq x \leq 1}(0.5, 0.04)$

(c) Triangular distribution $X \sim Trg(a, b, c) = Trg(0, 1, 0.3)$

(d) Triangular distribution $X \sim Trg(a, b, c) = Trg(0, 1, 0.7)$

(e) Triangular distribution $X \sim Trg(a, b, c) = Trg(0, 1, 0.0)$

Figure 5.5: Probability density function of demand share distribution

lization problem in our algorithm, as well as the LP relaxation of the problem and the original MIP to optimality. We implemented our algorithm in Julia 1.6.5. and enable parallelization on up to 32 threads.

We define the names of the problem instances by indicating the number of candidate facility locations (F), customers (D) and available capacity levels (C). For example, the problem instance F17D70C8 is a problem instance with 17 candidate facility locations, 70 customers and 8 available capacity levels. F17D70C8 also represents the real-world case of designing the hydrogen production infrastructure in Norway.

5.6.1 Comparison with the expected value problem

For instance F17D70C8, we calculate the solution to the deterministic expected value problem (EVP) and compare the results with the stochastic problem (SP) using 3, 50, and 100 scenarios. We study the performance of first-stage solutions on a reference sample with 1000 scenarios for each distribution. When solving the EVP, the different scenarios are replaced by a single scenario where all customers request their expected demand. The expected demand level is illustrated in Figure 5.6. Note that the normal and uniform distributions have identical expected demand and thus, identical EVPs, while each triangular distribution has different expected demand.

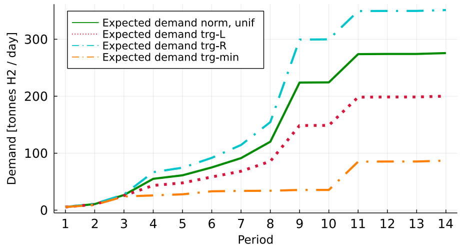


Figure 5.6: Expected demand level

The expected value of the EVP solution (EEV) is calculated by evaluating the first-stage solution from the EVP over the reference sample with 1000 scenarios. We further evaluate first-stage decisions from solving the SP using 3, 50, and 100 scenarios over the reference sample with 1000 scenarios. The objective value is then denoted RP. Note that problems with 3 scenarios are solved to optimality using Gurobi (the Lagrangian heuristic provides solutions with a proven optimality gap $< 2\%$). The problems with 50 and 100 scenarios are solved with a proven optimality gap $< 4\%$ using the Lagrangian heuristic, while Gurobi cannot find any feasible solution within three days of computing time. The value of the stochastic solution (VSS) given as $VSS = EEV - RP$ (see, e.g., Birge and Louveaux, 2011), provides a lower estimate of the true VSS since we do not solve the problem with 1000 scenarios to optimality.

		3 scen		50 scen		100 scen	
dist.	EEV [10^6]	RP [10^6]	VSS [%]	RP [10^6]	VSS [%]	RP [10^6]	VSS [%]
norm	37,531.9	3,134.3	91.65	2,958.6	92.12	2,945.5	92.15
unif	139,380.4	3,359.2	97.59	2,935.5	97.89	2,925.6	97.90
trg-L	2,742.5	2,726.8	0.57	2,695.0	1.73	2,692.1	1.84
trg-R	27,339,684.7	3,702.1	99.99	3,347.1	99.99	3,279.6	99.99
trg-min	2,565.8	2,502.0	2.49	2,318.6	9.63	2,274.3	11.36

Table 5.4: Out-of-sample evaluation for F17D70C8

Results in Table 5.4 show the EEV and the RP considering first-stage solutions obtained for 3, 50, and 100 scenarios. For each RP solution, we calculate the relative VSS to EEV. The EVP solution for normal, uniform and right-skewed triangular distribution is feasible only when using penalties for capacity excess for scenarios with low demand. Both left-skewed triangular distributions (trg-L, trg-min) have sufficiently low expected values so that the penalties for capacity excess are avoided. Simultaneously, capacity expansion in the second stage provides sufficient flexibility to always avoid penalties for demand shortfall. This also applies if only maritime demand is considered when determining the locations and initial capacity of the production facilities to be opened (i.e., the first-stage decisions).

When increasing the number of scenarios from 50 to 100, first-stage decisions based on a solution to SP with 100 scenarios lead to lower RP. The improvement in RP when using 100 scenarios instead of 50 is at least 5.9%, 14.4%, 1.2%, 10.6%, 7.9% for the normal, uniform and triangular distributions, respectively. Therefore, we focus on the results for 100 scenarios in our further evaluations.

In general, the installed capacity in symmetric and right-skewed EVPs is considerably higher than the highest installed capacity among SPs (see Figure 5.7) and therefore penalties for capacity excess apply in low-demand scenarios. The exceptions are the left-skewed distributions. The solution to EVP for trg-L installs slightly less capacity than the solution to SP in the first time periods while from period 9 onwards the installed capacity is higher. Note that in period 9, there is the most significant jump in the expected demand level for trg-L. The solution to SP leads to more conservative opening decisions to avoid low capacity utilization in scenarios where demand is realized below the expected value. Considering the distribution trg-min, the solution to SP installs more capacity to save expansion costs in scenarios where demand is realized above the expected value.

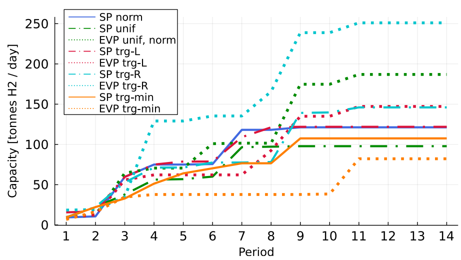


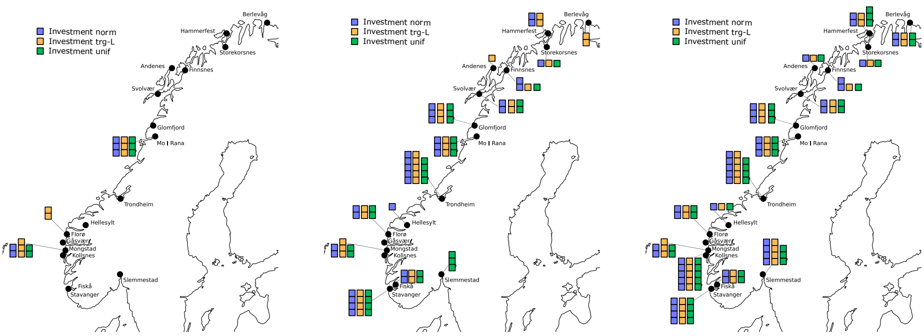
Figure 5.7: Installed capacity in the first stage (100 scenarios)

We can further observe (with exception of trg-min) that the installed capacity in SP is considerably lower than the expected demand while in the EVP, the installed capacity is close to the expected demand level. The reason is that capacity expansion is more expensive than opening a big facility right away. For a known demand level, the aim is to satisfy demand with very few expansions. Among the results for the SP, solutions for the uniform and the trg-min distributions lead to the lowest installed capacity. The uniform distribution is characterized by the highest variance among scenarios. Since the capacity level can be easily increased by expansion, the solution installs less capacity in the first stage to avoid low capacity utilization in scenarios with a low demand level.

When sampling multiple times, 3 scenarios are not always sufficient to avoid penalties even if the problem with 3 scenarios can be solved to optimality using Gurobi. In order to avoid penalties for capacity excess, at least one of the three scenarios has to be a scenario with a relatively low demand level which forces the solution to install less capacity. The solution to EEV for trg-min has shown that capacity expansion enables to increase the capacity, if necessary, and to avoid penalties for demand shortfall. Considering 50 and 100 scenarios, the probability of having a low-demand scenario in a sample is sufficiently large.

5.6.2 Solution structure

To analyze the opening decisions for different demand distributions, we study the structure of the first-stage decisions for instance F17D70C8, solved with 100 scenarios with a proven optimality gap $< 4\%$. We focus on the normal, left-skewed triangular, and uniform distributions as these are considered to reflect plausible demand scenarios for Norway. The geographical locations of opened facilities in different time periods are shown in Figure 5.8. We visualize the opened facilities in periods 1, 5, and 9, which allows us to analyze the main investment steps. Note that the solutions open the last new facility in period 9. From period 10 onwards, there are no additional first-stage opening decisions and demand increase is only compensated by capacity expansion.



(a) Opened facilities in time period 1

(b) Opened facilities in time period 5

(c) Opened facilities in time period 9

Figure 5.8: First-stage decisions: Investment in the SP

In the first period (see Figure 5.8a), the facilities are located in the middle of the southern and northern regions. These are strategic locations which can satisfy all customers without the necessity to open small local facilities. Surprisingly, the highest installed capacity is in the trg-L distribution, which is the distribution with the lowest expected demand. At the same time, this distribution is characterized by the lowest variance among scenarios. Therefore, the solution aims to install sufficient capacity to satisfy demand in more scenarios without expansion since opening right away a bigger facility is cheaper than expansion. These savings in investment costs compensate for higher production costs in low-demand scenarios with low capacity utilization. Figure 5.8b illustrates the opened facilities in period 5. We see that the solution opens most facilities when considering the trg-L distribution, while the fewest facilities are opened for the uniform distribution. The locations of the opened facilities are spread out along the entire coastline, irrespective of distribution. In period 9 (see Figure 5.8c), 16 out of 17 possible facilities are opened. Hydrogen production is characterized by economies of scale. However, high distribution costs dominate economies of scale in production and therefore the solution chooses to open many relatively small facilities.

For all distributions, the largest facilities are opened in Kollsnes and Trondheim. Kollsnes has a strategic position on the west coast of Norway as most of the maritime customers and road traffic customers in the southern part of Norway are located within 1000 km distribution distance. Trondheim is an important location as it is the only location in the northern region that can supply road traffic customers in the southern part of Norway. A facility in Trondheim can therefore exploit both lower production costs due to lower electricity prices and economies of scale in production due to supplying municipalities with high demand.

Most of the production is located in the southern region. Even if production is cheaper in the northern region, distribution costs are high. If the distance travelled from a facility in the southern region is about 470km shorter than from a facility in the northern region, it is favourable to use the facility located in the southern region. It would therefore be cheaper to supply all coastal customers south of Florø from local facilities rather than from Trondheim, even though the latter has cheaper production.

When considering a higher number of facilities, the number of opened facilities in the solution increases as well. Similar to the instance F17D70C8, due to high distribution costs, local production is preferred to centralized large-scale production in all instances. Further, most of the production is located in the southern region close to the customers. A facility located in the southern part of the northern region still plays plays a crucial role in supplying customers in the northern part of the southern region using the advantage of lower production prices. Further, a higher number of available capacities leads to lower objective. However, it does not lead to any structural changes in the infrastructure. The timing of the investment decisions is mostly affected by the demand curves and the fact that the production technology has minimum production requirements. Therefore, the opening of new facilities is in line with the demand increase. Since, the main characteristics of the solutions for larger instances are the same as we describe for instance F17D70C8, we have decided not to discuss them in detail.

5.6.3 Solution quality

To analyze the quality of our lower bound, we compare it with the optimal solution to the MIP and with the LP relaxation bound and calculate the optimality gap. Given the complexity of the problem, Gurobi can find optimal solutions and solve the LP relaxation only for a few instances with 3 scenarios even when allowing four days of computing time. In Table 5.5, we provide the results for two different samples and demand distributions for the instance F17D70C8, the instance F25D70C16 for the left-skewed triangular distribution and the instance F17D70C16 for the normal distribution since Gurobi can find an optimal solution within four days of computing time only for these instances. When increasing the size of the problem size, neither the MIP nor the LP relaxation can be solved within the time limit. Table 5.5 shows the objective value of the LP relaxation and the Lagrangian bound as well as the time needed to solve the LP relaxation to optimality. Since the MIP optimal solution is known, we calculate the optimality gaps. Our Lagrangian heuristic finds good lower bounds within three hours with an optimality gap only slightly higher than the one of the LP relaxation (about 0.5%). For larger instances, the LP relaxation cannot be solved within four days which highlights the importance of scalable methods such as our Lagrangian heuristic.

Instance	dist.	Gurobi LP relax. Obj.[10^6]	Gurobi LP relax. time[s]	LR bound 3h Obj.[10^6]	Opt. gap [%] LP relax.	Opt. gap [%] LR bound
F17D70C8	norm	3563.4	1290	3561.7	0.47	0.52
F17D70C8	trg-L	2789.5	1378	2789.3	0.44	0.45
F17D70C8	unif	3658.1	936	3657.3	0.46	0.48
F17D70C8	norm	2892.7	1813	2891.1	0.41	0.47
F17D70C8	trg-L	3007.6	3022	3007.3	0.55	0.56
F17D70C8	unif	3198.0	1200	3196.7	0.33	0.37
F17D70C16	norm	2757.1	77473	2756.8	0.50	0.51
F17D70C16	trg-L	-	-	3311.5	-	-
F17D70C16	unif	-	-	3082.0	-	-
F25D70C16	norm	-	-	2836.3	-	-
F25D70C16	trg-L	3268.1	305954	3267.0	0.45	0.49
F25D70C16	unif	-	-	3136.3	-	-

Table 5.5: Quality of the bounds for 3 scenarios

To discuss the performance of our algorithm and the quality of our solution, we show the results of our algorithm for 100 scenarios after 1 hour and after 5 hours

Instance	dist.	scen.	1 hour LB [10^6]	1 hour UB [10^6]	1 hour gap [%]	5 hours LB [10^6]	5 hours UB [10^6]	5 hours gap [%]	time to gap < 5%
F17D70C8	norm	10	2972.2	3061.9	2.93	2972.7	3058.7	2.81	218
F17D70C8	trg-L	10	2868.4	2938.8	2.40	2868.9	2935.6	2.27	229
F17D70C8	unif	10	3135.3	3256.5	3.72	3135.8	3224.0	2.74	246
F17D70C8	norm	25	3065.4	3166.8	3.20	3065.4	3166.8	3.20	747
F17D70C8	trg-L	25	2730.6	2834.9	3.68	2731.0	2826.8	3.39	1331
F17D70C8	unif	25	2986.9	3121.6	4.32	2987.6	3089.3	3.29	589
F17D70C8	norm	50	2931.6	3041.3	3.61	2931.9	3041.3	3.60	3474
F17D70C8	trg-L	50	2762.6	2892.9	4.51	2763.0	2863.6	3.51	2433
F17D70C8	unif	50	2944.8	3089.2	4.68	2945.2	3064.7	3.90	1273
F17D70C8	norm	100	2949.4	3095.2	4.71	2950.3	3063.5	3.70	3043
F17D70C8	trg-L	100	2766.4	2895.5	4.46	2767.0	2877.5	3.84	2610
F17D70C8	unif	100	2905.9	3061.4	5.08	2906.7	3024.4	3.89	5099
F17D70C16	norm	100	2929.6	3063.1	4.36	2930.1	3009.8	2.65	2937
F17D70C16	trg-L	100	2749.1	2849.6	3.53	2749.5	2843.7	3.31	2610
F17D70C16	unif	100	2882.0	2988.3	3.56	2882.4	2982.9	3.37	2712
F25D70C8	norm	100	2908.7	3058.2	4.89	2909.0	3039.8	4.28	3248
F25D70C8	trg-L	100	2729.0	2848.6	4.21	2729.7	2848.8	4.18	3021
F25D70C8	unif	100	2861.6	3041.0	5.90	2862.7	2989.6	4.25	8654
F25D70C16	norm	100	2892.3	3058.3	5.43	2892.9	2981.2	2.96	3813
F25D70C16	trg-L	100	2713.3	2853.9	4.93	2714.0	2812.4	3.50	3516
F25D70C16	unif	100	2844.5	2991.2	4.91	2845.1	2940.3	3.24	3298
F34D70C8	norm	100	2903.6	3086.4	5.92	2903.9	3044.1	4.60	8807
F34D70C8	trg-L	100	2724.7	2847.9	4.32	2725.4	2847.9	4.30	2980
F34D70C8	unif	100	2856.4	3054.8	6.49	2857.5	2995.8	4.61	8295
F34D70C16	norm	100	2888.6	3077.7	6.15	2889.2	3007.9	3.95	5185
F34D70C16	trg-L	100	2710.0	2841.9	4.64	2710.5	2810.9	3.57	3532
F34D70C16	unif	100	2841.2	3018.0	5.86	2841.6	2960.7	4.02	6267
F17D390C8	norm	100	2787.8	3066.9	9.10	2814.0	2902.6	3.05	4373
F17D390C8	trg-L	100	2421.5	2618.8	7.53	2450.7	2548.5	3.83	4819
F17D390C8	unif	100	2669.1	2817.6	5.27	2695.7	2802.1	3.80	4068
F17D390C16	norm	100	2767.5	2997.2	7.66	2795.7	2867.9	2.52	5128
F17D390C16	trg-L	100	2396.8	2637.9	9.14	2433.5	2500.7	2.69	4474
F17D390C16	unif	100	2654.7	2919.9	9.08	2677.5	2754.7	2.80	3845
F25D390C8	norm	100	2727.5	3139.2	13.11	2779.0	2878.8	3.47	7727
F25D390C8	trg-L	100	2362.7	2734.2	13.58	2418.4	2515.8	3.87	10339
F25D390C8	unif	100	2611.9	2962.0	11.82	2662.5	2780.5	4.24	8531
F25D390C16	norm	100	2703.3	3027.8	10.72	2763.6	2857.0	3.27	6159
F25D390C16	trg-L	100	2337.4	2723.4	14.18	2403.6	2491.8	3.54	6326
F25D390C16	unif	100	2596.3	2997.7	13.39	2646.3	2734.1	3.21	5129
F34D390C8	norm	100	2697.1	3101.6	13.04	2772.2	2903.0	4.50	9855
F34D390C8	trg-L	100	2372.2	2631.1	9.84	2412.5	2526.2	4.50	6848
F34D390C8	unif	100	2581.4	3006.3	14.13	2627.1	2760.6	4.84	8951
F34D390C16	norm	100	2695.1	3009.3	10.44	2779.0	2878.9	3.47	7857
F34D390C16	trg-L	100	2315.3	2778.4	16.67	2400.5	2509.8	4.35	10645
F34D390C16	unif	100	2517.1	3098.3	18.76	2642.6	2761.1	4.29	8118

Table 5.6: Computational results

of computing time in Table 5.6. We show the lower and upper bound, as well as the resulting gap and provide the computing time required to achieve a gap lower than 5%. For instance F17D70C8, we also show the results of our algorithm for

10, 25, 50, and 100 scenarios to assess the scalability of our algorithm when the number of scenarios increases.

Our Lagrangian heuristic finds good feasible solutions for all tested instances within a time limit of one hour. When increasing the time limit to 5 hours, we also observe a slight improvement in the lower bound. However, the main improvement is due to finding better feasible solutions with lower objective function values. After 5 hours, a solution with a proven optimality gap $< 5\%$ can be found for all tested instances.

The results show that for instances with a small number of scenarios, we find solutions with lower optimality gaps than for instances with a higher number of scenarios, since more iterations are performed and it is easier to find a first-stage solution that avoids penalties for all scenarios. However, the difference in solution quality between 50 scenarios and 100 scenarios is minimal.

Surprisingly, when increasing the problem size from 8 to 16 capacities, the resulting optimality gap tends to decrease, just as the run time needed to achieve a gap $< 5\%$. With 16 capacities, it is easier to find a suitable capacity level for the required production quantities than with 8. Therefore, our upper bound heuristic finds good solutions with low optimality gaps already in early iterations.

Since we allow for parallelization on up to 32 threads when calculating the lower bound, we further observe that increasing the number of candidate facility locations has a relatively low impact on the quality of our solution. However, with 34 candidate locations, we can see that the resulting gap increases as well as the time needed to achieve a gap $< 5\%$, since the number of iterations performed during the computing time decreases.

The instances with 390 customers are characterized by relatively long computing times to achieve an optimality gap $< 5\%$. The time needed to update the Lagrangian multipliers increases as the size of the problem (5.29)–(5.33) depends on the number of customers and scenarios. In later iterations, updating the multipliers takes approximately 70% of the time needed for one iteration. When increasing the time limit, we see a considerable improvement in the upper bound. However, the lower bound improves in average by 2%.

5.7 Conclusion

We have studied the problem of locating hydrogen production facilities in Norway under demand uncertainty. We have formulated our problem as a two-stage stochastic multi-period facility location and capacity expansion problem considering minimum production requirements. The state-of-the-art commercial solver Gurobi can solve only the smallest instances with a low number of scenarios. Since the out-of-sample performance can be improved considerably when increasing the number of scenarios, we present a solution method based on Lagrangian relaxation to solve larger problems with a higher number of scenarios. With our algorithm, we find high-quality solutions for all tested instances within five hours computing time.

Results for small test instances indicate that our algorithm provides good lower bounds. Thus, for future work, the improvement potential lies within the upper bound heuristic. We further observed that the box-step method is a limiting factor for instances with a large number of customers, as the time needed to update the Lagrangian multipliers increases considerably. If shorter computing times are needed, exploring different methods or a combination of methods for the calculation of the Lagrangian dual may be a promising direction.

When solving our facility location model for the problem of locating hydrogen production in Norway, we see that due to high distribution costs, the solution chooses to open facilities at most of the candidate locations. Furthermore, most of the production is located in the southern part of Norway, since high distribution costs dominate the lower production costs in the northern part of Norway. The facility in Trondheim is therefore characterized by high opening capacity as it has low production costs and many road traffic customers in the southern part of Norway are within the distance limit. However, the demand scenarios used in our analysis are characterized by a large degree of uncertainty. We show that different distribution types do not have a large impact on size and location of the opened facilities. Still, more precise input data, in particular for future hydrogen demand, may provide a better basis for generating the scenario tree for our problem. Additional research efforts should therefore be dedicated to estimating future hydrogen demand, but this is outside the scope of the analysis in this paper.

The model can be further extended by including the choice of production technology. Together with including uncertainty in investment and production costs, the model might also be used to capture uncertainty in technology development. This is subject to future research.

Bibliography

Aglen, T. M. and Hofstad, A. (2022). Designing the hydrogen supply chain for maritime transportation. Master's thesis, Department of Industrial Economics and Technology Management, NTNU, Trondheim, Norway.

Almansoori, A. and Shah, N. (2012). Design and operation of a stochastic hydrogen supply chain network under demand uncertainty. *International Journal of Hydrogen Energy*, 37(5):3965–3977.

Amiri, A. (1997). Solution procedures for the service system design problem. *Computers & Operations Research*, 24(1):49–60.

Andrenacci, S., Yejung, C., Raka, Y., Talic, B., and Colmenares-Rausseo, L. (2022). *Electrolysers towards EU MAWP 2023 targets and beyond*. Zenodo.

Angulo, G., Ahmed, S., and Dey, S. S. (2016). Improving the integer l-shaped method. *INFORMS Journal on Computing*, 28(3):483–499.

Balachandran, V. and Jain, S. (1976). Optimal facility location under random demand with general cost structure. *Naval Research Logistics Quarterly*, 23(3):421–436.

Birge, J. R. and Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media, New York, 2 edition.

Christensen, T. R. L. and Klose, A. (2021). A fast exact method for the capacitated facility location problem with differentiable convex production costs. *European Journal of Operational Research*, 292(3):855–868.

Correia, I. and Captivo, M. E. (2003). A Lagrangean heuristic for a modular capacitated location problem. *Annals of Operations Research*, 122(1):141–161.

Correia, I. and Melo, T. (2021). Integrated facility location and capacity planning under uncertainty. *Computational and Applied Mathematics*, 40(5):1–36.

Correia, I. and Saldanha da Gama, F. (2019). Facility location under uncertainty. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, pages 185–213. Springer.

Crew, B. (2022). Solving the energy crisis. *Nature*, 609(7926):S1–S1.

Danebergs, J. and Aarskog, F. G. (2020). Future compressed hydrogen infrastructure for the domestic maritime sector. IFE/E-2020/006, Halden, Norway.

Dayhim, M., Jafari, M. A., and Mazurek, M. (2014). Planning sustainable hydrogen supply chain infrastructure with uncertain demand. *International Journal of Hydrogen Energy*, 39(13):6789–6801.

DNV GL (2019). Produksjon og bruk av hydrogen i Norge. Rapport 2019-0039, Oslo, Norway, (in Norwegian).

Govindan, K., Fattahi, M., and Keyvanshokooh, E. (2017). Supply chain network design under uncertainty: A comprehensive review and future research directions. *European Journal of Operational Research*, 263(1):108–141.

Holmberg, K. (1994). Solving the staircase cost facility location problem with decomposition and piecewise linearization. *European Journal of Operational Research*, 75(1):41–61.

IEA (2022). Global hydrogen review 2022. *International Energy Agency*, pages 1–284.

Jakobsen, D. and Åtland, V. (2016). Concepts for large scale hydrogen production. Master's thesis, Department of Energy and Process Engineering, NTNU, Trondheim, Norway.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2015). Dynamic facility location with generalized modular capacities. *Transportation Science*, 49(3):484–499.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2016). Solving a dynamic facility location problem with partial closing and reopening. *Computers & Operations Research*, 67:143–154.

Jena, S. D., Cordeau, J.-F., and Gendron, B. (2017). Lagrangian heuristics for large-scale dynamic facility location with generalized modular capacities. *INFORMS Journal on Computing*, 29(3):388–404.

Kim, J., Lee, Y., and Moon, I. (2008). Optimization of a hydrogen supply chain under demand uncertainty. *International Journal of Hydrogen Energy*, 33(18):4715–4729.

Kleywegt, A. J., Shapiro, A., and Homem-de Mello, T. (2001). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

Li, L., Manier, H., and Manier, M.-A. (2019). Hydrogen supply chain network design: An optimization-oriented review. *Renewable and Sustainable Energy Reviews*, 103:342–360.

Li, X. and Zhang, K. (2018). A sample average approximation approach for supply chain network design with facility disruptions. *Computers & Industrial Engineering*, 126:243–251.

Lucas, C., MirHassani, S., Mitra, G., and Poojari, C. (2001). An application of lagrangian relaxation to a capacity planning problem under uncertainty. *Journal of the Operational Research Society*, 52(11):1256–1266.

Marsten, R. E., Hogan, W. W., and Blankenship, J. W. (1975). The boxstep method for large-scale optimization. *Operations Research*, 23(3):389–405.

Melo, M. T., Nickel, S., and Saldanha da Gama, F. (2009). Facility location and supply chain management–a review. *European Journal of Operational Research*, 196(2):401–412.

NEL Hydrogen (2018). Nel hydrogen electrokysers. `http://img-admin.exponews.com.au.s3.amazonaws.com/exhibitors/e/nel-electrolysers-brochure-2018-pd-0600-0125-web_18041145.pdf/`. last accessed 01.08.2022.

Nickel, S. and Saldanha da Gama, F. (2019). Multi-period facility location. In Laporte, G., Nickel, S., and Saldanha da Gama, F., editors, *Location Science*, pages 303–326. Springer, Cham.

Nord Pool AS (2022). Day-ahead prices. `https://www.nordpoolgroup.com/en/Market-data1/Dayahead/Area-Prices/ALL1/Yearly/?view=table/`. last accessed 01.08.2022.

Nunes, P., Oliveira, F., Hamacher, S., and Almansoori, A. (2015). Design of a hydrogen supply chain with uncertainty. *International Journal of Hydrogen Energy*, 40(46):16408–16418.

NVE (2021). Langsiktig kraftmarkedsanalyse 2021 – 2040. `https://publikasjoner.nve.no/rapport/2021/rapport2021_29.pdf`. last accessed 01.08.2022, (in Norwegian).

Ocean Hyway Cluster (2020a). 2030 hydrogen demand in the Norwegian domestic maritime sector. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Ocean Hyway Cluster (2020b). Interactive map - potential maritime hydrogen in Norway. OHC HyInfra project, Workpackage C: Mapping future hydrogen demand.

Owen, S. H. and Daskin, M. S. (1998). Strategic facility location: A review. *European Journal of Operational Research*, 111(3):423–447.

Regjeringen (2019). Handlingsplan for gronn skipsfart. `https://www.regjeringen.no/contentassets/2ccd2f4e14d44bc88c93ac4effe78b2f/handlingsplan-for-gronn-skipsfart.pdf`. last accessed 09.03.2022,(in Norwegian).

Santoso, T., Ahmed, S., Goetschalckx, M., and Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115.

Sauvey, C., Melo, T., and Correia, I. (2020). Heuristics for a multi-period facility location problem with delayed demand satisfaction. *Computers & Industrial Engineering*, 139:106171.

Schütz, P., Stougie, L., and Tomasgard, A. (2008). Stochastic facility location with general long-run costs and convex short-run costs. *Computers & Operations Research*, 35(9):2988–3000.

Schütz, P., Tomasgard, A., and Ahmed, S. (2009). Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199(2):409–419.

Sherali, H. D. and Zhu, X. (2006). On solving discrete two-stage stochastic programs having mixed-integer first-and second-stage variables. *Mathematical Programming*, 108(2):597–616.

Shulman, A. (1991). An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations Research*, 39(3):423–436.

Snyder, L. V. (2006). Facility location under uncertainty: a review. *IIE Transactions*, 38(7):547–564.

Štádlerová, Š., Aglen, T. M., Hofstad, A., and Schütz, P. (2022a). Locating hydrogen production in norway under uncertainty. In Ramalhinho, H., De Armas, J., and Voß, S., editors, *Computational Logistics*, volume 13557, pages 306–321. Springer, Cham.

Štádlerová, Š. and Schütz, P. (2021). Designing the hydrogen supply chain for maritime transportation in Norway. In Mes, M., Lalla-Ruiz, E., and Voß, S., editors, *Computational Logistics*, volume 13004, pages 36–50. Springer, Cham.

Štádlerová, Š., Schütz, P., and Tomasgard, A. (2022b). Multi-period facility location and capacity expansion with modular capacities and convex short-term costs. Working paper, Department of Industrial Economics and Technology Management, NTNU, Norway.

Statistics Norway (2018). Statistics Norway: 12579: Road traffic volumes. `https://www.ssb.no/en/statbank/table/12579/`. Online; accessed 02.11.2021.

United Nations (2015). Paris agreement. In *Adoption of the Paris agreement*. Framework Convention on Climate Change, FCCC/CP/2015/L.9/Rev.1.

Van den Broek, J., Schütz, P., Stougie, L., and Tomasgard, A. (2006). Location of slaughterhouses under economies of scale. *European Journal of Operational Research*, 175(2):740–750.

ISBN 978-82-326-7326-1 (printed ver.)
ISBN 978-82-326-7325-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Norwegian University of
Science and Technology