CIRRELT

CIRRELT-2024-09

Solving Multi-Stage Stochastic Facility Location Problems with Modular Capacity Adjustments

Šárka Štádlerová
Peter Schütz
Sanjay Dominik Jena

April 2023

Bureau de Montréal
Université de Montréal
C.P. 6128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

Solving Multi-Stage Stochastic Facility Location Problems with Modular Capacity Adjustments

Šárka Štádlerová[1,*], Peter Schütz[1], Sanjay Dominik Jena[2]

1. Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology
2. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Analytics, Operations, and Information Technologies Department (AOTI), School of Management, Université du Québec à Montréal

Abstract. We consider a multi-stage stochastic facility location problem with modular capacity adjustments, minimizing the expected costs of allocating uncertain customer demand. We present a general multi-stage mixed-integer programming formulation that allows for multiple facility expansions, reductions, and closing of existing facilities. Given the complexity of this planning problem, we present a solution method based on Lagrangian relaxation, followed by the solution of a restricted model to further improve the solution quality. The computational results show that our solution method provides high-quality solutions within reasonable computing times. We further compare the value of a multi-stage stochastic solution to the solution of a deterministic rolling horizon problem and discuss situations when solving a multi-stage problem is particularly beneficial.

Keywords: Multi-stage stochastic facility location, Lagrangian relaxation, Capacity adjustments.

Acknowledgements. This work 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. The work of the third author was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada under grant 2017-05224.

---

* Corresponding author: sarka.stadlerova@ntnu.no

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2024

1 Introduction

Establishing hydrogen infrastructure and switching to hydrogen fuels in the transportation sector is an important step towards meeting the ambitious objectives of the Paris Agreement (United Nations, 2015). A diversified energy mix is an important milestone for achieving energy stability (IEA, 2022). Green hydrogen, i.e., hydrogen produced without emitting CO_2, must be produced through water electrolysis (EL) using renewable electricity. With 99% of electricity being produced from renewable sources, 88% from hydropower and 11% from wind parks (Ministry of Petroleum and Energy, 2023), Norway has a good starting position for the production of green hydrogen. However, both electric batteries and ammonia are competing zero-emission alternatives. The future energy mix, and hence future hydrogen demand, is still uncertain (DNV GL, 2019). A flexible infrastructure allowing for capacity expansion as well as reduction to account for the uncertainty in future hydrogen demand, becomes crucial to satisfy customers and provide a reliable hydrogen supply.

Motivated by the planning context of locating hydrogen facilities in Norway, we here study a location problem that minimizes the expected total costs while satisfying customers demand as best as possible under three distinguishing characteristics: modular capacity adjustments, allowing to adjust production capacities to changing demand trends; a piece-wise linear objective-function that models the piece-wise linear short-term costs more realistically, specific to each capacity level; and the multi-stage stochastic nature of the problem, allowing for the adjustment of decisions throughout the planning horizon when more precise information on the uncertain parameters becomes available. While such characteristics has been tackled separately in the existing literature, its combination, required by the here given application context, has not yet been considered.

Our contributions are threefold. First, we present a general multi-stage stochastic facility location model allowing for multiple capacity adjustments during the planning horizon. The proposed model formulation allows for a large variety of different multi-stage configurations. It further extends the deterministic generalized facility location model by Jena et al. (2015) by considering uncertainty in future demand within a multi-stage setting. We further model specific piece-wise linear short-term costs dependent on installed capacity and the geographical location of the facility. Second, we present a solution method based on Lagrangian relaxation to find high-quality solutions to real-world size multi-stage facility location problems with capacity adjustments within reasonable computational time. To the best of our knowledge, our work is the first to apply Lagrangian relaxation to relax the demand constraint in the context of multi-stage stochastic facility location problems. The Lagrangian dual is solved by a hybrid strategy to update the Lagrangian multipliers more efficiently: we first use a cutting planes method with box constraints that allows for fast convergence in early iterations; we then switch to a steepest descent method due to its fast computing times in subsequent iterations. Finally, we evaluate the performance of the proposed solution method on a variety of problems instances derived from an industrial case study

The remainder of this paper is structured as follows. Section 2 reviews the related literature. We then define the planning problem and present a mathematical model in Section 3. Section 4 focuses on the solution method. The case study considered in our computational experiments is introduced in Section 5. Computational results and conclusions are discussed in Sections 6 and 7, respectively.

2 Related Literature

In the following, we review literature related to the three key characteristics of the here considered location problem: modular capacity adjustments, piece-wiese linear objective functions and parameter uncertainty in the context of multi-stage planning.

Capacity adjustments can be applied during the planning horizon in response to time-varying parameters, such as demand or costs. As a result, depending on the application context, capacity expansion, reduction or relocation may be necessary. Deterministic facility location problems with capacity adjustments have received a lot of attention in the literature. An overview on such works can be found in the reviews of Melo et al. (2009), Nickel and Saldanha-da Gama (2019), Alarcon-Gerbier and Buscher (2022).

Capacity adjustments have initially be modeled by closing existing facilities and opening new ones at the same location (see, e.g., Shulman, 1991; Dias et al., 2007). More recent literature uses modular capacities, where capacity adjustments can be modelled as a change between available capacity levels. Correia et al.

(2013), Cortinhal et al. (2015), Šťádlerová et al. (2022), and Šťádlerová et al. (2024) allow for the expansion of capacity over time. Problems that allow for capacity expansion and reduction, along with facility closing and reopening have been discussed by Antunes and Peeters (2001), Jena et al. (2015), Jena et al. (2017), and Becker et al. (2019). Melo et al. (2006), Jena et al. (2016), and Allman and Zhang (2020) study capacity modification by relocation of modules.

A particular advantage of modular capacities is that they allow for the modelling of economies of scale in investment and production costs, since such costs can be defined for each capacity level (see, e.g., Correia and Captivo, 2003). To model production costs more in detail, a specific production cost function depending on capacity utilization can be associated with each capacity level which also enables to consider economies of scale in production and minimum production limits for each capacity level resulting from technology specifications in industrial processes (see, e.g., Šťádlerová et al., 2022, 2024). Such detailed cost functions are considered in this work within the context of general modular capacity adjustments and a multi-stage planning context under demand uncertainty.

The latter, parameter uncertainty, is a natural extension of deterministic formulations. Within the paradigm of stochastic programming (Birge and Louveaux, 2011), the literature has mainly focused on two-stage stochastic facility location problems that have been formulated as single-period problems. For a review on facility location problems under uncertainty, we refer to Owen and Daskin (1998), Snyder (2006), Govindan et al. (2017), Correia and Saldanha-da Gama (2019). Multi-period two-stage stochastic facility location problems with capacity expansion as a second-stage decision are relatively new in the literature. Correia and Melo (2021), and Šťádlerová et al. (2022, 2023) study a version of the problem, where facility location and its initial capacity are first-stage decisions while capacity expansion, together with demand allocation, are second-stage decisions depending on the realization of uncertain demand parameters.

Formulating the planning problem with multiple decision stages (Ortiz-Astorquiza et al., 2018) allows for the revision of capacity decisions as more information about uncertain parameters is revealed, which may lead to considerable cost savings. However, the computational complexity increases compared to the two-stage formulation. Indeed, previous studies on multi-stage stochastic facility location and capacity expansion problems show that due to the complexity of the problem and a high number of binary variables, only problems of limited size with relatively few scenarios can be solved (see, e.g., Ahmed et al., 2003; Huang and Ahmed, 2009; Singh et al., 2009). One of the first studies on multi-stage stochastic facility location problems with capacity expansion is provided by Ahmed et al. (2003). The authors present a model with continuous capacity expansion variables and only require the overall capacity to be higher than demand. They use Lagrangian relaxation to obtain a valid bound, by relaxing the non-anticipativity constraints resulting in a problem that can be decomposed in scenarios. They then construct an integer solution for each scenario, and finally re-enforce the non-anticipativity constraints and construct a feasible solution. The solution is further improved using a branch and bound algorithm. Due to the computational complexity, the problem is solved only for small instances. Huang and Ahmed (2009) formulate a multi-stage facility location problem with capacity expansion as a recourse decision. They further show the advantage of a multi-stage formulation over a two-stage formulation by calculating the value of multi-stage stochastic programming (VMSS). For a review on multi-stage stochastic optimization problems, we refer to Bakker et al. (2020).

Multi-stage stochastic location problems with modular capacities have not yet been considered in the literature. The present work aims at filling this gap, additionally considering detailed capacity-level dependent piece-wise linear cost-functions and proposing a Lagrangian heuristic to find high-quality solutions to real-world sized problem instances.

3 Problem formulation

We first provide a detailed definition of the here considered planning problem in Section 3.1. A general mixed-integer programming (MIP) formulation is then given in Section 3.2.

3.1 Problem definition

We present a location and capacity investment planning problem considering time-varying and uncertain demand. The problem naturally has a multi-stage structure, since the exact information about future demand becomes available gradually during the planning horizon. At the beginning of each stage, when new information becomes available, facilities can be opened, and existing facilities can expand capacity, reduce capacity, or be closed in response to the expected or observed demand level. Considering the time-varying uncertain demand, facilities need to be opened and modified so that demand is satisfied. Otherwise, penalty costs for unsatisfied demand must be paid.

Let $\mathcal{I}$ represent the set of candidate facility locations and $\mathcal{K} = \{-1, 0, 0', 1, 2, \ldots, \overline{k}\}$ the set of available capacity levels. Capacity levels may have different interpretations. Specifically, capacity level 0 indicates that the location currently does not have any facility, but is available for opening a new facility in the future. Capacity level -1 indicates that a previously existing facility has been permanently shut down, and reopening in the future is not allowed. Capacity level $0'$ is used if the facility has been closed, but a new facility can be opened later again at that location. Capacity levels $1, 2, \ldots, \overline{k}$ indicate an existing facility open at that capacity production level. We further define a capacity subset $\mathcal{K}^0 \subset \mathcal{K} = \{-1, 0, 0'\}$, containing capacity levels without production capacity, and a capacity subset $\mathcal{K}^+ \subset \mathcal{K} = \{1, 2, \ldots, \overline{k}\}$, containing production capacity levels. As such, $\mathcal{K} = \mathcal{K}^+ \cup \mathcal{K}^0$ contains all non-zero capacity levels as well as capacity levels without production capacity. Note that $\mathcal{K}^+ \cap \mathcal{K}^0 = \emptyset$. Capacity set $\mathcal{L}^-(i, t, k_1) \subseteq \mathcal{K}$ defines all capacity levels to which the capacity of facility $i \in \mathcal{I}$, currently open at capacity level k_1, can be changed to at the beginning of time period $t \in \mathcal{T}$. Similarly, capacity set $\mathcal{L}^+(i, t, k_2) \subseteq \mathcal{K}$ contains all capacity levels from which the capacity may be changed to level k_2 at facility $i \in \mathcal{I}$ at the beginning of time period $t \in \mathcal{T}$.

Customer locations are given by the set $\mathcal{J}$. Let $\mathcal{S}$ be the set of scenarios s, where s is the representation of future uncertain demand level. Each customer has a specific demand D^s_{jt} for time period $t \in \mathcal{T}$ in scenario $s \in \mathcal{S}$. Customer demand can be satisfied by one or more facilities, which, in the facility location literature, is often referred to as multi-sourcing. Parameter L_{ij} is set to value 1 if facility $i \in \mathcal{I}$ can serve customer $j \in \mathcal{J}$, and to value 0 otherwise. Unit distribution costs T_{ij} are calculated based on the distance between facility $i \in \mathcal{I}$ and customer $j \in \mathcal{J}$. If demand cannot be satisfied, penalty costs M^D are paid for each unsatisfied demand unit d^s_{jt}. Such penalty costs may, for example, reflect additional costs for importing the product.

Investment and production costs are subject to economies of scale and they depend both on the size of the installed capacity and the location of the facility. The costs for changing a capacity level from $k_1 \in \mathcal{K}$ to capacity level $k_2 \in \mathcal{L}^-(i, t, k_1)$ at the beginning of time period $t \in \mathcal{T}$ is given by $C_{ik_1k_2t}$. Parameter Φ_{ik} takes value 1 if the initial capacity level at the beginning of the planning horizon is equal to $k \in \mathcal{K}$ at location $i \in \mathcal{I}$. Costs for changing capacity level $C_{ik_1k_2t}$ represent long-term costs and they are separated from short-term production costs.

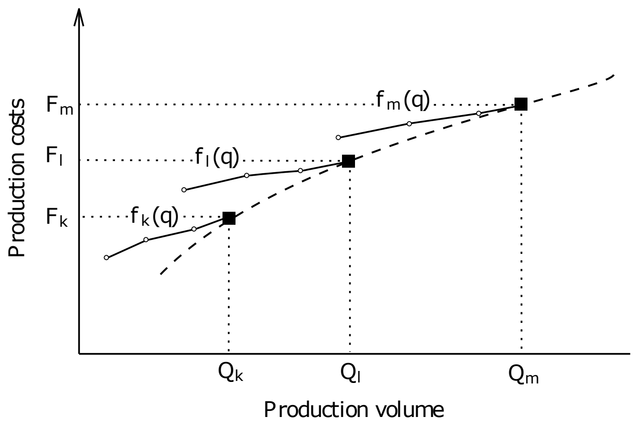


Figure 1: Short-term production costs function

For each capacity level, a specific piece-wise linear short-term production cost function $f_k(q)$ defines both the cost and the feasible production quantities for the installed capacity as illustrated in Figure 1. At each break-point, production quantities and corresponding costs are defined. The lowest break-point of the short-term production costs function represents the lowest production quantities at a given capacity level while the highest break-point $\overline{B}_k$ corresponds to the installed capacity and thus to the upper production limit at capacity level k. The short-term production costs for given location $i \in \mathcal{I}$, production capacity level $k \in \mathcal{K}^+$, break-point $b \in \mathcal{B}_k$ and time period $t \in \mathcal{T}$ are denoted F_{ikbt}. Using a linear combination of breakpoints, arbitrary quantities between the minimum and maximum limit can be achieved. While Šťádlerová et al. (2024) are limited to strictly convex short-term production costs, we here consider a general cost function that can take any shape. Having the assumption of convex short-term production costs, the optimal solution always combines two consecutive breakpoints to get the required production quantities. In the case of a general shape of the short-term cost function, the piece-wise linear function is modelled using a special ordered set of type 2 (Schütz et al., 2008; Williams, 2013).

When considering minimum production requirements, penalty costs for overproduction M^Q apply for each overproduction unit q^s_{it} if demanded quantities are below the level of minimum production requirements. Similar to the penalty costs for unsatisfied demand, penalty costs for overproduction may represent costs for exporting excess production.

We here opt to model demand uncertainty by means of a multi-stage scenario-tree that represents alternative demand scenarios $s \in \mathcal{S}$. The scenario-tree is defined by a set of nodes $\mathcal{N}$. Hence, we define each scenario as the entire set of demands for all locations throughout the entire planning horizon. A scenario therefore consists of a sequence of nodes $n \in \mathcal{N}$. Let $\mathcal{S}(n)$ be the set of scenarios passing through node $n \in \mathcal{N}$ and $t_n \in \mathcal{T}$ the time period associated with node n. In other words, all scenarios in $\mathcal{S}(n)$ cannot yet be distinguished from each other and they still have the exact same demand realizations in the node $n \in \mathcal{N}$ and the time period t_n.

Our model can deal a large variety of multi-stage decision structures, including an arbitrary number of decision stages and different number of time-periods within each stage. The scenario-tree structure of a simple three-stage problem is depicted in Figure 2. All scenarios passing through the node $n \in \mathcal{N}$ in time period t_n are characterized with equal decision structure in t_n. In Figure 2, the scenarios $1, 2$, and 3 lead to equal decisions until time period 4, since they are passing through the same nodes and cannot be distinguished from each other.

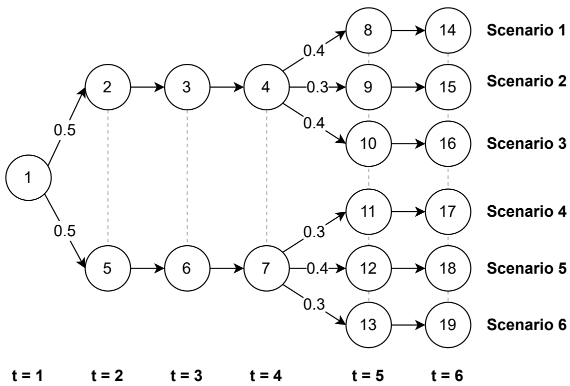


Figure 2: Graphical representation of the three-stage scenario-tree

3.2 Mathematical formulation

We first introduce the decision variables used to model the planning problem defined above as a mixed-integer linear program:

$y^s_{ik_1k_2t}$	1, if facility at location $i \in \mathcal{I}$ changes capacity level from $k_1 \in \mathcal{K}$ to capacity level $k_2 \in \mathcal{L}^-(i,t,k_1)$ at the beginning of period $t \in \mathcal{T}$ for scenario $s \in \mathcal{S}$; 0 otherwise;
μ^s_{ikbt}	Weight of breakpoint $b \in \mathcal{B}_k$ at location $i \in \mathcal{I}$ for capacity level $k \in \mathcal{K}^+$ in period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$;
x^s_{ijkt}	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 $s \in \mathcal{S}$;
d^s_{jt}	Demand shortfall: the amount of unsatisfied demand at customer location $j \in \mathcal{J}$ in period $t \in \mathcal{T}$ for scenario $s \in \mathcal{S}$;
q^s_{it}	Overproduction: the amount of overproduction at facility location $i \in \mathcal{I}$ in period $t \in \mathcal{T}$ for scenario $s \in \mathcal{S}$.

The resulting multi-stage stochastic model can be written as follows:

$$\min \sum_{s\in\mathcal{S}} p^s \Bigg[\sum_{i\in\mathcal{I}} \sum_{t\in\mathcal{T}} \sum_{k_1\in\mathcal{K}} \sum_{k_2\in\mathcal{L}^-(i,t,k_1)} C_{ik_1k_2t} y^s_{ik_1k_2t} + \sum_{b\in\mathcal{B}_k} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}^+} \sum_{t\in\mathcal{T}} F_{ikbt}\mu^s_{ikbt} + \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}^+} \sum_{t\in\mathcal{T}} T_{ij} x^s_{ijkt} \sum_{j\in\mathcal{J}} \sum_{t\in\mathcal{T}} M^D d^s_{jt} + \sum_{i\in\mathcal{I}} \sum_{t\in\mathcal{T}} M^Q q^s_{it} \Bigg], \tag{1}$$

subject to:

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}^+} x^s_{ijkt} + d^s_{jt} = D^s_{jt}, \qquad j\in\mathcal{J}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{2}$$

$$\sum_{k_1\in\mathcal{K}} \sum_{k_2\in\mathcal{L}^-(i,t,k_1)} y^s_{ik_1k_2t} = 1, \qquad i\in\mathcal{I}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{3}$$

$$\sum_{k_2\in\mathcal{L}^-(i,t,k)} y^s_{ikk_2t} = \Phi_{ik}, \qquad i\in\mathcal{I}, t=1, k_1\in\mathcal{K}, s\in\mathcal{S}, \tag{4}$$

$$\sum_{k_1\in\mathcal{L}^+(i,t-1,k)} y^s_{ik_1k(t-1)} = \sum_{k_2\in\mathcal{L}^-(i,t,k)} y^s_{ikk_2t}, \qquad i\in\mathcal{I}, t\in\mathcal{T}\setminus\{1\}, k\in\mathcal{K}, s\in\mathcal{S}, \tag{5}$$

$$\sum_{b\in\mathcal{B}_{k_2}} \mu^s_{ik_2bt} = \sum_{k_1\in\mathcal{L}^+(i,t,k_2)} y^s_{ik_1k_2t}, \qquad i\in\mathcal{I}, t\in\mathcal{T}, k_2\in\mathcal{K}^+, s\in\mathcal{S}, \tag{6}$$

$$\sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}^+} x^s_{ijkt} + q^s_{it} = \sum_{b\in\mathcal{B}_k} \sum_{k\in\mathcal{K}^+} Q_{bk}\mu^s_{bikt}, \qquad i\in\mathcal{I}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{7}$$

$$x^s_{ijk_2t} \le L_{ij} D^s_{jt} \sum_{k_1\in\mathcal{L}^+(i,t-1,k_2)} y^s_{ik_1k_2t}, \qquad i\in\mathcal{I}, t\in\mathcal{T}, j\in\mathcal{J}, k_2\in\mathcal{K}^+, s\in\mathcal{S}(n), \tag{8}$$

$$\frac{1}{|\mathcal{S}(n)|} \sum_{s'\in\mathcal{S}(n)} y^{s'}_{ik_1k_2t_n} = y^s_{ik_1k_2t_n}, \tag{9}$$

$i\in\mathcal{I}, n\in\mathcal{N}, k_1\in\mathcal{K}, k_2\in\mathcal{L}^-(i,t,k_1), n\in\mathcal{N}, s\in\mathcal{S}(n),$

$$y^s_{ik_1k_2t} \in \{0,1\}, \qquad i\in\mathcal{I}, t\in\mathcal{T}, k_1\in\mathcal{K}, k_2\in\mathcal{L}^-(i,t,k_1), s\in\mathcal{S}, \tag{10}$$

$$x^s_{ijkt} \ge 0, \qquad i\in\mathcal{I}, j\in\mathcal{J}, k\in\mathcal{K}^+, t\in\mathcal{T}, s\in\mathcal{S}, \tag{11}$$

$$\mu^s_{ikbt} \ge 0, \qquad b\in\mathcal{B}_k, i\in\mathcal{I}, k\in\mathcal{K}^+, t\in\mathcal{T}, s\in\mathcal{S}, \tag{12}$$

$$q^s_{it} \ge 0, \qquad i\in\mathcal{J}, t\in\mathcal{T}, s\in\mathcal{S}, \tag{13}$$

$$d^s_{jt} \ge 0, \qquad j\in\mathcal{J}, t\in\mathcal{T}, s\in\mathcal{S}. \tag{14}$$

Objective (1) minimizes the expected sum of investment, capacity adjustments, production, and distribution costs, as well as the penalty costs for unsatisfied demand and overproduction.

Equations (2) ensure that demand is either satisfied or accounted for as demand shortfall. Constraints (3) require that only one capacity level can be selected for each facility location. Constraints (4) initialize the capacity levels at the beginning of the planning horizon. Constraints (5) are the flow conservation constraints linking the capacity change variables throughout consecutive time periods. Equalities (6) link the selected capacity level to the appropriate short-term cost function. Constraints (7) ensure that the entire production is either distributed to customers or that overproduction is penalized. If minimum production requirements have to be considered, the quantity Q_{bk} given by the smallest breakpoint b will be strictly positive. For problems without minimum production requirements, the quantity belonging to the smallest breakpoint is zero. Constraints (8) limit which facility can satisfy which customer and link the distribution variable to the operated capacity level. These constraints are formulated in the form of strong inequalities, which are known to provide stronger linear relaxation bounds than aggregated linking constraints (see, e.g., Jena et al., 2017). Constraints (9) require that scenarios $s \in \mathcal{S}(n)$ passing through the node $n \in \mathcal{N}$ in time period t_n are characterized by equal decisions in time period t_n. Constraints (10) – (14) are the non-negativity and binary requirements.

4 Solution method

We solve the multi-stage stochastic problem (1) – (14) using a solution approach based on Lagrangian relaxation. Lagrangian relaxation has been widely used to solve deterministic multi-period facility location problems (see, e.g., Shulman, 1991; Jena et al., 2016, 2017; Šťádlerová et al., 2024). Šťádlerová et al. (2023) have applied Lagrangian relaxation to a two-stage stochastic multi-period facility location problem with capacity expansion, showing that Lagrangian relaxation provides tight bounds.

4.1 Lagrangian relaxation

Relaxing the demand constraints has been a popular choice in the literature, given that this tends to result in a decomposition by facility locations (see, e.g., Shulman, 1991; Jena et al., 2016, 2017; Šťádlerová et al., 2024). We follow this approach, relaxing the demand constraints (2), which are indeed the only linking constraints connecting the different facility locations. Let λ^s_{jt} be the Lagrangian multipliers belonging to demand constraints (2). The resulting Lagrangian relaxation is formulated as follows:

$$
\begin{aligned}
LR(\boldsymbol{\lambda}) = \min \sum_{s\in\mathcal{S}} p^s \Bigg[& \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}}\sum_{k_1\in\mathcal{K}}\sum_{k_2\in\mathcal{L}^-(i,t,k_1)} C_{ik_1k_2t} y^s_{ik_1k_2t} + \\
& \sum_{b\in\mathcal{B}_k}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}^+}\sum_{t\in\mathcal{T}} F_{ikbt}\mu^s_{ikbt} + \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}} M^Q q^s_{it} + \\
& \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}^+}\sum_{t\in\mathcal{T}} (T_{ij} - \lambda^s_{jt}) x^s_{ijkt} + \\
& \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}} (M^D - \lambda^s_{jt}) d^s_{jt} + \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}} \lambda^s_{jt} D^s_{jt} \Bigg],
\end{aligned} \tag{15}
$$

subject to Constraints (3) – (14).

For given 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. It can further be shown that any optimal solution to $LR(\boldsymbol{\lambda})$ will set d^s_{jt} to 0 (Šťádlerová et al., 2023). We therefore omit both terms when solving the Lagrangian relaxation.

Since all remaining constraints are defined for each facility location $i \in \mathcal{I}$, the problem can be decomposed and solved by facility location. We hence define $LR(\boldsymbol{\lambda}) = \sum_{i\in\mathcal{I}} g_i(\lambda) + \sum_{s\in\mathcal{S}}\sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}} p^s\lambda_{jt}D^s_{jt}$, where $g_i(\boldsymbol{\lambda})$ is the optimal objective function value of the Lagrangian subproblem for facility location i:

$$
\begin{aligned}
g_i(\boldsymbol{\lambda}) = \min \sum_{s\in\mathcal{S}} p^s \Bigg[& \sum_{t\in\mathcal{T}}\sum_{k_1\in\mathcal{K}}\sum_{k_2\in\mathcal{L}^-(i,t,k_1)} C_{ik_1k_2t} y^s_{ik_1k_2t} + \\
& \sum_{b\in\mathcal{B}_k}\sum_{k\in\mathcal{K}^+}\sum_{t\in\mathcal{T}} F_{ikbt}\mu^s_{ikbt} + \sum_{t\in\mathcal{T}} M^Q q^s_{it} + \\
& \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}^+}\sum_{t\in\mathcal{T}} (T_{ij} - \lambda^s_{jt}) x^s_{ijkt} \Bigg],
\end{aligned} \tag{16}
$$

subject to those Constraints (3) – (14) that are defined for facility i.

4.1.1 Solving the Lagrangian subproblem

In order to obtain optimal opening and capacity adjustment decisions together with optimal demand allocation for a given facility in all time periods and scenarios, we formulate the Lagrangian subproblem for a single candidate facility location as an expected shortest-path problem and solve it using dynamic programming. The shortest path formulation is well known in the literature for deterministic problems (Shulman, 1991; Jena et al., 2016, 2017; Šťádlerová et al., 2024). The expected shortest path formulation for a two-stage stochastic facility location problem with a limited number of expansions is studied in Šťádlerová et al. (2023). We generalize this approach from two stages to multiple stages and formulate the expected shortest path for a multi-stage problem allowing for capacity expansion and capacity reduction multiple times during the planning horizon. Similar to Šťádlerová et al. (2023), the expected shortest path has a nested structure. Therefore, the non-anticipativity constraints in the multi-stage structure are considered. Hence, for a given facility, the result of the shortest path problem provides optimal opening and capacity adjustment schedule for the entire planning horizon in all scenarios and this solution does not violate the non-anticipativity constraints of the original problem (9).

In order to solve the shortest path problem, the demand allocation costs for given facility location $i \in \mathcal{I}$, capacity $k \in \mathcal{K}$, time period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$ need to be computed. The demand allocation problem costs can be calculated by means of continuous knapsack as explained in the next section.

4.1.2 Continuous knapsack problem

The problem of optimal demand allocation for given facility $i \in \mathcal{I}$, capacity $k \in \mathcal{K}$, time period $t \in \mathcal{T}$ and scenario $s \in \mathcal{S}$ can be formulated as a continuous knapsack problem with piecewise linear costs (Amiri, 1997). For given capacity, the demand allocation problem is specific for each scenario $s \in \mathcal{S}$ and can therefore be solved independently for each scenario. As proposed by Šťádlerová et al. (2023), the problem can be written as:

$$K^s_{ikt}(\boldsymbol{\lambda}) = \min \sum_{b\in\mathcal{B}_k} F_{ikbt}\mu^s_{ikbt} + M^Q q^s_{it} + \sum_{j\in\mathcal{J}} (T_{ij} - \lambda^s_{jt}) x^s_{ijkt}, \tag{17}$$

subject to:

$$x^s_{ijkt} \le L_{ij} D^s_{jt}, \qquad j \in \mathcal{J}, \tag{18}$$

$$\sum_{j\in\mathcal{J}} x^s_{ijkt} + q^s_{it} = \sum_{b\in\mathcal{B}_k} Q_{bl}\mu^s_{ikbt}, \tag{19}$$

$$\sum_{b\in\mathcal{B}_k} \mu^s_{ikbt} = 1, \tag{20}$$

$$q^s_{it} \ge 0, \tag{21}$$

$$x^s_{ijkt} \ge 0, \qquad j \in \mathcal{J}, \tag{22}$$

$$\mu^s_{ikbt} \ge 0, \qquad b \in \mathcal{B}_k. \tag{23}$$

4.2 Solution of the Lagrangian dual problem

By solving the relaxed problem (15), subject to Constraints (3) – (14), for given multipliers λ^s_{jt}, we obtain a lower bound on the Objective (1). To find the best possible lower bound, we have to solve the Lagrangian dual problem $LD = \max_\lambda LR(\boldsymbol{\lambda})$, which aims at identifying the optimal multipliers $\boldsymbol{\lambda}$. We first use the cutting planes method (Section 4.2.1) with box constraints before switching to the subgradient method (Section 4.2.2), since the cutting planes method becomes time-consuming in later iterations.

4.2.1 Cutting planes method

The Lagrangian dual problem can be solved by several methods. Šťádlerová et al. (2023) use a cutting planes method with box constraints, an adoption of its deterministic version (see, e.g., Marsten et al., 1975; Schütz et al., 2009; Šťádlerová et al., 2024) that converges quickly and requires a reasonable computational effort during the first iterations. However, in later iterations, the computational effort increases as the problem contains more cutting planes. As a result, updating the multipliers takes about 99% of the computing time when solving for the lower bound, while solving the relaxed problem accounts for only 1%. In contrast, the classical subgradient method (see, e.g., Polyak, 1969) has a slower convergence, but the computational effort remains relatively stable throughout the iterations. We therefore use the cutting planes method with box constraints for the first 200 iterations to converge quickly, and then further improve the lower bound by computing further iterations with the subgradient method.

Adopting the cutting planes method with box constraints from Šťádlerová et al. (2023), we calculate the elements of the subgradient matrix $\boldsymbol{\nabla}$ as $\nabla^{ms}_{jt} = D^s_{jt} - \sum_{i\in\mathcal{I}} x^{sm}_{ijt}$ at each iteration m and for each scenario s, where x^{sm}_{ijt} is the solution to the Lagrangian subproblem obtained for variables $\boldsymbol{x}$ in iteration m. We then define $L^m = LR(\boldsymbol{\lambda}^m) - \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}} p^s \lambda^{ms}_{jt}\nabla^{ms}_{jt}$. To find the updated multipliers, we solve the following linear optimization problem:

$$\max \phi \tag{24}$$

$$\phi \leq L^g + \sum_{j\in\mathcal{J}}\sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}} p^s \nabla^{gs}_{jt}\lambda^{m+1,s}_{jt}, \qquad g = 1, ..., m, \tag{25}$$

$$\lambda^{ms}_{jt} - \Delta^{ms}_{jt} \leq \lambda^{m+1,s}_{jt} \leq \lambda^{ms}_{jt} + \Delta^{ms}_{jt}, \qquad j \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}, \tag{26}$$

$$\phi \in \mathbb{R}, \lambda^{m+1,s}_{jt} \in \mathbb{R}. \tag{27}$$

Constraints (26) are the box constraints and control the changes of the Lagrangian multipliers. Marsten et al. (1975) show that the ideal box size depends on the problem and cost structure and the convergence can be further improved if the box size is adjusted dynamically. We choose an initial box size of 100 and decrease the box size using the coefficient 0.85 each time the sign of of ∇^{ms}_{jt} changes.

4.2.2 Subgradient method

While the subgradient method has a rather slow convergence (Jena et al., 2017), it is fast to compute and can find new multipliers that improve the lower bound. We therefore switch to the subgradient method after 200 iterations to reduce the computational effort. The approach to update the multipliers is adopted from Van den Broek et al. (2006) and Jena et al. (2017). In each iteration m, the step size θ^m is calculated as:

$$\theta^m = \eta^m \frac{UB - LR(\boldsymbol{\lambda})^m}{\|\nabla^{ms}_{jt}\|^2},$$

where UB represents the best upper bound found so far and η^m is a scalar. The new Lagrangian multipliers $\lambda^{m+1,s}_{jt}$ for the $(m+1)th$ iteration are obtained as:

$$\lambda^{m+1,s}_{jt} = \lambda^{ms}_{jt} + \theta^m \nabla^{ms}_{jt}, j \in \mathcal{J}, t \in \mathcal{T}, s \in \mathcal{S}.$$

We follow the suggestion by Shulman (1991) to choose the initial scalar value $\eta^m = 2$ and then divide it by 2 each time the lower bound does not improve for n consecutive iterations. Based on initial tests, we choose $n = 5$.

4.3 Calculating feasible upper bounds

In general, the solution to the Lagrangian relaxation does not provide a feasible solution to the original problem. We present a heuristic in Section 4.3.1 that generates a feasible solution from the solutions to the Lagrangian subproblems and therefore provides an upper bound to the original problem. In Section 4.3.2, we present a restricted MIP approach to further improve this solution.

4.3.1 Lagrangian heuristic

Since the solution to the relaxed problem is obtained separately for each facility, it is usually not feasible for the original problem. We therefore use an heuristic approach to find a feasible solution, and thus an upper bound, to the original problem. The main idea of our heuristic is depicted in Figure 3.

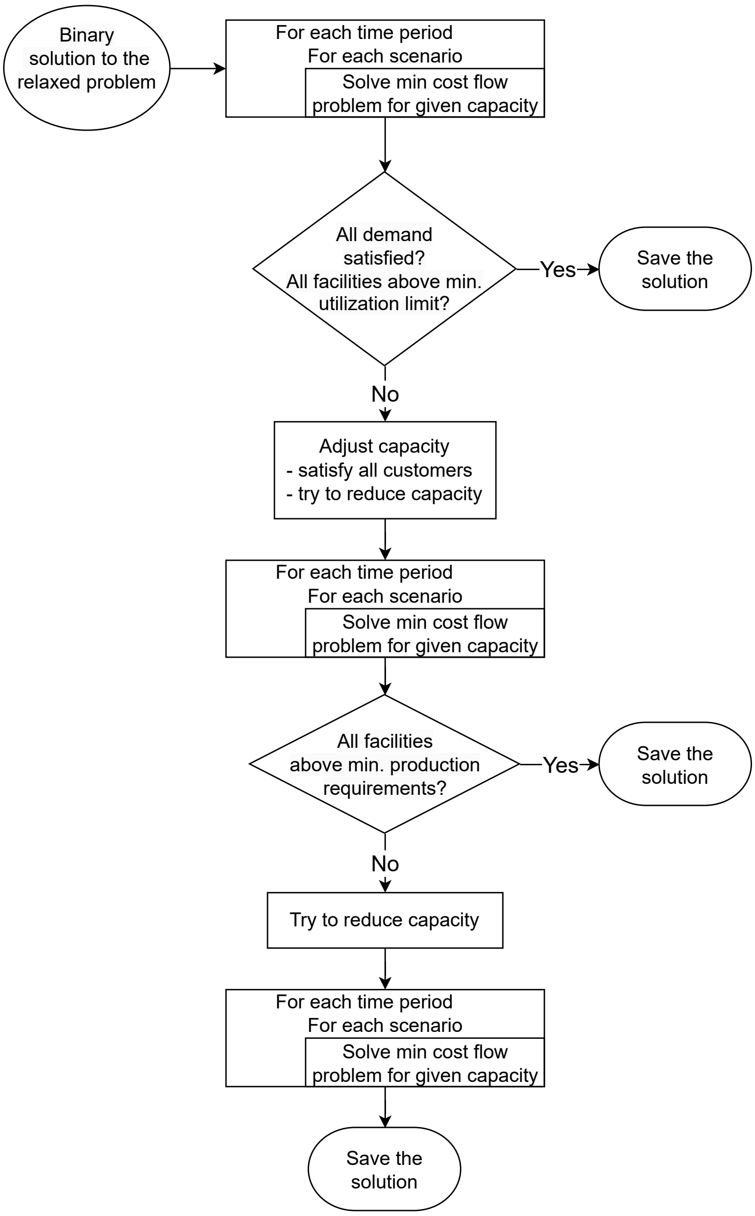


Figure 3: Heuristic for finding an upper bound

We first find an allocation of demand that ensures that demand is at most satisfied once. We use the locations and capacities proposed by the solutions to the Lagrangian subproblems, and find the optimal demand allocation for each time period and scenario by solving a minimum cost flow problem. During this initial allocation of demand, we relax minimum production requirements. To ensure relatively complete recourse, we introduce an artificial facility with sufficiently large capacity (in case the available capacity is insufficient) and an artificial customer with sufficiently large demand (in case of overproduction).

The solution to the minimum cost flow problem provides valuable information about capacity utilization at given facilities, as well as about demand satisfaction. We use this information to modify the capacities in an attempt to find an optimal flow without the artificial facility or customer. Note that all capacity changes in period t_n in scenario $s \in S(n)$ must be replicated for all other scenarios in the set $S(n)$.

Using the artificial facility or artificial customer implies penalty costs. In an attempt to improve the planning solutions, we aim to satisfy all customer demand from real facilities. If customer demand is unsatisfied (i.e., served from the artificial facility) or facilities overproduce (i.e., they deliver to the artificial customer), we can distinguish between two strategies:

- Delivering from the artificial facility: Increase capacity to satisfy the demand of all customers in all time periods and scenarios from existing facilities. We iteratively increase the size of the facility that can satisfy most of the unsatisfied customers and reallocate customers from the artificial facility. In case of a tie, we choose the facility with the lowest sum of distribution costs to customers.
- Delivering to the artificial customer: Reduce capacity of all facilities with overproduction. We calculate the lowest capacity that is needed to satisfy the flow to real customers and decrease the capacity accordingly.

After changing the capacity, we resolve the minimum cost flow problem to find a new demand allocation using the updated facility capacities. We repeat this procedure until all customer demand is satisfied. However, some facilities may operate below the minimum utilization limit. In an iterative process over all time periods, we verify the minimum production requirements for each facility in each scenario. In case a facility operates below the minimum utilization limit in time period t_n and at least one scenario $s \in S(n)$, we verify whether or not the facility has been opened in period t_n: If time period t_n is the opening period of the facility, we remove the facility for time period t_n and all scenarios in set $S(n)$. As a result, the heuristic postpones the opening to $t_n + 1$. If the facility has been opened prior to period t_n, we reduce its capacity using a similar logic as for removing the facility. In case we remove or reduce the capacity of a facility, the minimum costs flow problem is solved once again.

4.3.2 Restricted MIP

To further improve the quality of the solution, we use a restricted MIP approach (R-MIP) (see, e.g., Jena et al., 2017; Šťádlerová et al., 2024). To this end, we store the n best solutions found with our upper bound heuristic and fix all binary variables y to 1, if $\frac{1}{n}\sum y \geq 0.8$. With these variables being fixed, we use Gurobi to solve the original problem (1) – (14). Note that the optimal solution to the restricted MIP provides an upper bound to the optimal solution of the original problem. Based on preliminary testing, we set $n = 5$.

5 Case study

We consider the case of locating hydrogen production for the Norwegian transport sector. The production facilities are to be located throughout Norway to satisfy customer demand during the planning horizon. Our case study builds on Šťádlerová et al. (2023) that uses uncertain hydrogen demand (monotonically increasing over time) from maritime passenger transportation, road-based transportation and offshore operations in the oil and gas sector. We use the same data for technology, costs and demand, but change how the demand scenarios are generated.

We first present how we generate the demand scenarios for the multi-stage model presented in this paper. For the sake of completeness, we also provide an overview of the costs used in our case study. Finally, we summarize the characteristics of the resulting set of problem instances.

5.1 Customer locations and demand

We consider up to 70 customer locations and a planning horizon represented by 14 time periods. For each customer location j, the demand D^s_{jt} at time period t and scenario s is composed of a deterministic component and two independent stochastic components: The deterministic component represents demand from maritime passenger transportation, where public contracts specify operations and schedules over long periods of time. We therefore consider maritime demand as deterministic and present in all demand scenarios. The two stochastic components represent demand from land-based transportation and the offshore sector. These sectors are more difficult to predict and thus scenario-dependent.

In all instances, the overall demand is non-decreasing over time for each scenario. However, there might be some structural changes in customer locations over time. We therefore generate two different types of scenario trees, each of which represents different tendencies of demand evolution over time:

- Increasing: Demand is non-decreasing at all customer locations and for all scenarios, similar to Štádlerová et al. (2023).
- Mixed: During the first time periods $t = 1, \ldots, t' - 1$, demand is non-decreasing at all customer locations and in all scenarios. Starting in period t', demand may decrease at some customer locations. This decrease is compensated by a demand increase at some other customer locations in the particular scenario. Overall demand is still non-decreasing in all scenarios.

Additional details on how maximum aggregated daily demand develops over the planning horizon are given in A.

Demand scenarios for the "Increasing" scenario trees are generated as follows: Demand quantities D^s_{jt} for customer location j at time period t in scenario $s \in \mathcal{S}$ are computed as $D^s_{jt} = D^D_{jt} + w^1_t(s)D^{S1}_{jt} + w^2_t(s)D^{S2}_{jt}$, where D^D_{jt} represents the deterministic component. D^{S1}_{jt} and D^{S2}_{jt} are the maximum demands of the stochastic demand components given by demand forecasts for the land-based sector (DNV GL, 2019) and the offshore sector (Ocean Hyway Cluster, 2020; Aglen and Hofstad, 2022), respectively. Coefficients $w^1_t(s)$ and $w^2_t(s)$ are the scaling factors of the first and second stochastic demand component, respectively. The scaling factors of the two stochastic components are computed as follows: For each scenario $s \in \mathcal{S}$, we first uniformly sample $W^i(s)$ between 0 and 1, representing the maximum share of D^i_{jt}, $i \in \{1, 2\}$. In the first period, $t = 1$, the scaling factor $w^i_1(s)$ is a random number between 0 and $W^i(s)$. For all subsequent time periods, we then independently sample the scaling factor $w^i_t(s)$ such that $w^i_{t-1}(s) \leq w^i_t(s) \leq W^i(s)$.

To generate scenarios for the "Mixed" scenario trees, we first create an "Increasing" scenario tree. We then randomly choose 30% of the scenarios. In each of the selected scenarios s', we randomly select 30% of the customers. Demand for time periods $1, \ldots, 7$ for those customers remain unchanged, while demand starting at time period $t' = 8$ is modified as follows. For each of the customer locations, we sample a random number $V^i_j(s')$, $i \in \{1, 2\}$, between 0 and $w^i_{t'-1}(s')$. We then sample the reduced share $v^i_{jt}(s')$ for each period $t \geq t'$ such that $V^i_j(s') \leq v^i_{jt}(s') \leq w^i_{t-1}(s')$. Demand for this location is then updated as $D^{s'}_{jt} = D^D_{jt} + v^1_{jt}(s')D^{S1}_{jt} + v^2_{jt}(s')D^{S2}_{jt}$, $\forall t \geq t'$. Finally, we sample a customer location whose demand is increased by this demand reduction.

We generate scenario trees with 6, 12, 30, 90, 120, and 300 scenarios for both types of trees.

5.2 Candidate facility locations and costs

We consider up to 17 candidate locations. The investment and capacity adjustment costs in our case study are independent of facility location. We approximate the long-term cost function by 8 modular capacity levels, each with its specific short-term production cost function. The short-term production cost function is location-dependent. We distinguish between two regions: the southern region (S) with high production costs and the northern region (N) with low production costs. Investment costs and production costs at 100% capacity utilization are given in Table 4 in B. The candidate locations and capacity intervals are taken from Štádlerová et al. (2023), whereas the investment and production costs are derived from Jakobsen and Åtland (2016).

We use alkaline electrolysis as hydrogen production technology. The production range is 15%–100% (NEL Hydrogen, 2018). We approximate the short-term production cost function by a piece-wise linear function with 3 line pieces and 4 breakpoints at 15%, 50%, 80%, and 100% of the installed capacity level.

The costs for adjusting the capacity of facility i from capacity level $k_1 > 0$ to capacity level $k_2 > 0$ are calculated as $C_{ik_1k_2t} = (C_{i0k_2t} - C_{i0k_1t}) \cdot 1.15$ if $k_1 \leq k_2$, and as $C_{ik_1k_2t} = (C_{i0k_2t} - C_{i0k_1t}) \cdot 0.5$ if $k_1 > k_2$. The costs of permanently shutting down a facility (i.e., $k_2 = -1$) are equal to 50% of the investment costs. If facility has been closed but a new one can be opened later at the location (i.e., $k_2 = 0'$), we consider costs equal to 60% of the investment costs.

Hydrogen is distributed by trucks. Distribution costs per kilometre and kilogram of the product are provided in Table 5 in B. We set distribution limit equal to 1000km. Hence, distribution over distances > 1000 km is not allowed.

We define penalty costs for demand shortfall, M^Q, as well as for overproduction, M^D. The focus of the case study is on independent production and the import or export of hydrogen is not desirable. We therefore set high penalties of 10^6 €/kg for both shortfall and overproduction.

5.3 Instances

Based on the hydrogen case study, we generate 12 main problem instances, all having 17 candidate facility locations, 70 customers, and 8 capacity levels. However, they differ in the number of scenarios and the type of scenario tree. Based on these main problem instances, we further derive smaller instances with only 7 or 10 facilities and 10, 20, or 50 customers, using a subset of the original candidate facility and customer locations. For testing purposes, we run the smaller instances and instances with higher capacity adjustment costs and/or limited distribution range. We also test the performance of the Lagrangian method on the case study instances with 16 and 20 capacities where we also consider lower distribution costs (divided by 10).

We denote the instances by indicating the number of candidate facility locations (F), customer demand points (D), available capacity levels (C) and scenarios (S). We further distinguish samples with only increasing demand at all customer locations (I) and samples with decreasing as well as increasing demand (M).

6 Computational results

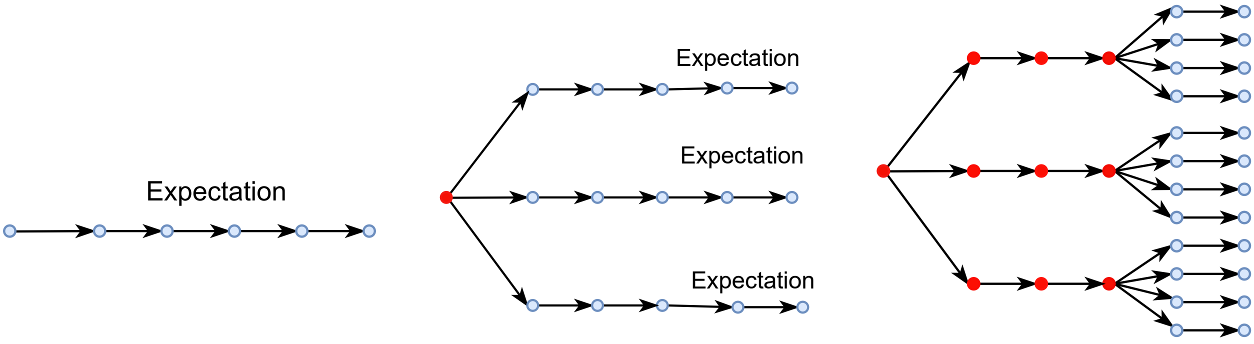
In this section, we present and discuss our computational results. All experiments have been carried out on a computer with two 3.6 GHz Intel Xeon Gold 6244 CPUs and 384 GB RAM, running Linux kernel 3.10. We use Gurobi Optimizer 10.0 for solving the deterministic rolling horizon problem and smaller instances of the three-stage problem to optimality as well as for solving the restricted MIP problem. We further use HiGHS solver (Huangfu and Hall, 2018) to solve the minimum cost flow problem to optimality. We implement our algorithm in Julia 1.8.2 and enable parallelization on up to 32 threads.

We first analyze the value of the 3-stage stochastic programming solution as opposed to solving the deterministic expected value problem in Section 6.1. We then explore the difficulty of solving the problem with a general-purpose solver and by means of the proposed Lagrangian heuristic in Section 6.2.

6.1 Value of multi-stage stochastic programming

The value of the stochastic solution (VSS) measures the quality of the stochastic solution compared to a solution derived from a simpler deterministic problem, often the expected value problem (Birge and Louveaux, 2011). While VSS for two-stage stochastic programming problems is well-defined, different approaches exist for estimating VSS in multi-stage stochastic programming problems (see e.g., Escudero et al., 2007; Nickel et al., 2012). In this paper, we adopt the approach used by Nickel et al. (2012) to determine the value of the multi-stage stochastic solution (VMSS). We first solve the deterministic expected value problem for the entire planning horizon. We then fix the first-stage decisions and solve another expected value problem for each of the possible realizations in the second stage. We continue to fix decisions and to solve expected value problems for all subsequent stages until we have solved the expected value problem for the last stage. This procedure is illustrated in Figure 4 for a three-stage problem, where blue nodes indicate decisions for time periods that lie ahead, while red nodes indicate fixed decisions. The VMSS is then given (for a minimization problem) as the difference between the objective function values of the sequence of deterministic expected value problems, v^D, and the objective function value of the multi-stage stochastic solution, v^S, i.e. $VMSS = v^D - v^S$. The relative value of the multi-stage stochastic solution (RVMSS) can be expressed as $RVMSS = \frac{v^D - v^S}{v^D}$.

To analyze the value of multi-stage stochastic solution, we use a set of instances that varies in number of candidate locations, capacity adjustment costs and distribution limits. Table 1 summarizes the results for such instances considering samples with decreasing as well as increasing demand. All instances are solved with Gurobi. The deterministic problems are all solved to optimality, whereas we provide the results after 24 hours for the three-stage stochastic problem (3-stage). Given that the optimal solution to the the multi-stage problem is generally unknown, we use the objective function value of the best known stochastic solution to estimate the RVMSS, providing a lower bound on the true RVMSS.



(a) First stage (b) Second stage (c) Third stage

Figure 4: Solving the sequence of deterministic expected value problems. Each node represents one time period. Red nodes correspond to fixed decisions, whereas blue nodes correspond to future decisions.

Instance	Adj. costs	Dist. limit	Objective [10^6] Det.	3-stage	Gap [%] 3-stage	RVMSS [%]
F10D10C8S12	1	1	2194.2	2179.6	0.30	0.67
F10D10C8S12	2	1	2307.4	2212.7	0.45	4.10
F10D10C8S12	5	1	2392.9	2212.7	0.26	7.53
F10D10C8S12	10	1	2477.8	2213.9	0.26	10.65
F17D10C8S12	1	1	2095.6	2095.1	0.25	0.02
F17D10C8S12	2	1	2201.2	2119.0	1.10	3.73
F17D10C8S12	5	1	2209.6	2117.7	1.08	4.16
F17D10C8S12	10	1	2209.3	2117.1	1.03	4.17
F17D10C8S12	1	0.5	2115.8	2113.2	0.34	0.12
F17D10C8S12	2	0.5	2252.6	2152.2	1.64	4.46
F17D10C8S12	5	0.5	2397.5	2154.4	1.62	10.14
F17D10C8S12	10	0.5	2636	2154.7	1.55	18.26

Table 1: Comparison of solutions to the deterministic and three-stage formulations

Note that "Adj costs" and "Dist limit" in Table 1 refer to the scaling factor used, where a scaling factor of 1 corresponds to the original adjustment costs in the case study and a distribution limit of 1000km, respectively. The RVMSS varies quite a lot across the different instances. For instances with relatively low costs of adjusting capacities, the RVMSS is less than 1%. In case of high capacity adjustment costs (scaling factor 5 and 10) however, the RVMSS increases to more than 18%, indicating that revising decisions throughout the planning horizon can be highly beneficial. We also see that a reduced distribution limit tends to increase the RVMSS, while an increased number of candidate locations lowers the RVMSS. Both adjustment costs, distribution limits and number of candidate locations impact the flexibility inherent in the production system as they affect to which degree decisions made in earlier stages can be adapted to observed demand. Solving the stochastic problem, and thus explicitly accounting for uncertainty, becomes more important with reduced flexibility in the system, i.e., high adjustment costs, lower distribution limits, and/or a small number of facility locations. These results are in line with the observations made by Schütz and Tomasgard (2011) who use a two-stage stochastic programming model for operational supply chain planning problem and report a low VSS when the underlying production system has a high level of flexibility.

6.2 Solution quality

To assess the performance of our Lagrangian-based method, we solve instances of different sizes and with different types of scenario trees and compare our results to the results obtained from Gurobi. The results are shown in Table 2. We report the optimality gap from the Lagrangian relaxation after a maximum of 6 hours computing time ("LR 6h"). As the Lagrangian relaxation converges faster for smaller instances, we also provide the actual run time for the Lagrangian relaxation. Once the Lagrangian relaxation terminates, we proceed with the R-MIP approach, allowing an additional 6 hours run time to further improve the solution. We only report a gap for the R-MIP in Table 2 if the R-MIP finds a new best solution. Gurobi is limited to a run time of 24 hours.

To analyze the performance of our Lagrangian-based method, we show the results after 6 hours run time ("LR 6h"). For smaller instances, the Lagrangian multipliers stop changing within 6 hours. If the Lagrangian multipliers have not changed for 5 consecutive iterations, we terminate the algorithm and report the time to convergence ("Converge"). After 6 hours or if terminating the Lagrangian relaxation due to convergence, we run the R-MIP to further improve the solution allowing additional 6 hours for Gurobi's branch-and-bound method. If the R-MIP does not find a new best solution, we report no optimality gap for the R-MIP. The results from our approach are compared to the optimality gap obtained by running Gurobi for 24 hours. All optimality gaps are calculated using the best known bound (see column "Best bound"). In general, Gurobi provides the best known bound for small instances, whereas our Lagrangian relaxation provides best known bounds for the larger instances. Note that Gurobi fails to solve the linear relaxation for all instances where we report the best known bound from the Lagrangrian relaxation. For smaller instances, the Lagrangian bound deviates only on average of 0.26% from the best Gurobi bound, indicating that the bounds provided by the Lagrangian heuristic are likely to be tight.

Gurobi can find feasible solutions for only 23 instances, even when provided a large time limit of 24 hours. The two smallest instances are solved to optimality within 3 hours. For instances with 90 or more scenarios, Gurobi cannot find a feasible solution. In contrast, the Lagrangian-based method provides solutions of good quality already after 6 hours for all instances. The deviations to the best known bounds tend to be between 2% and 4%, while the proven optimality gap is consistently smaller than 5% for all tested instances. In combination with the R-MIP, we can reduce the optimality gap for 37 of the 61 instances. In general, the R-MIP works well for smaller instances where the average improvement is 1.36%, while for large instances, the R-MIP may fail to find new improving feasible solutions.

We also study how scaling the adjustment costs and the distribution limit affect the quality of the solutions. We therefore solve the instances shown in Table 1 using our Lagrangian-based approach. In addition, we solve instances with 16 and 20 capacity levels instead of 8. The results for all of these instances are shown in Table 3. Note that all instances use a scenario tree with decreasing as well as increasing demand demand and that the distribution costs in instances marked with an asterisk are reduced using the factor of 0.1.

When solving instances F(·)D10C8S12 with higher adjustment costs and stricter distribution limits with our Lagrangian-based approach, the optimality gaps show a tendency to increase. This applies both to the Lagrangian relaxation and the R-MIP. The time for the Lagrangian relaxation to converge however, appears less affected by the change in adjustment costs and distribution limits. Increasing the number of candidate locations from 10 to 17 has a small impact on the time to convergence, increasing the computing time by only 1.5 minutes on average. The R-MIP approach can generally improve the optimality gap, but does not terminate within the runtime limit of 6 hours for any but the smallest instance without scaling of adjustment costs or distribution limit.

Using Gurobi to solve these instances, we first see a slightly different pattern in the behaviour of the optimality gap: With a small increase in adjustment costs, the gap starts to increase, before the gap starts decreasing with even higher adjustment costs. This might be due to the problem becoming more complex when the adjustment costs increase as correcting a decision becomes more expensive in subsequent stages. Once adjustment cost are sufficiently high, certain solution structures seem to be dominated, facilitating the solution of the planning problem.

An increase in the number capacity levels reduces the overall optimality gaps produced by our Lagrangian relaxation. Having more capacity levels for the facilities allows the model to better match observed demand with installed capacities, thus reducing the need for costly capacity adjustments. The gap for the instances with reduced distribution costs is a bit higher, resulting from a more complex problem instance. The added complexity is due to the increased relative importance of determining the right locations and capacities for the facilities. The disadvantage of the added capacity levels is that due to the size of the problem, our approach does no longer converge within the time limit of 6 hours. However, neither Gurobi nor the R-MIP can solve such large problem instances within their respective time limits.

Instance	Scen. tree	Gur 24h Gap [%]	LR 6h Gap [%]	Best bound	Converge Time [s]	R-MIP Gap [%]	R-MIP Time [s]
F10D10C8S6	I	0	1.78	Gur	301	0.51	1056
F10D10C8S6	M	0	2.63	Gur	325	0.95	328
F10D10C8S12	I	0.10	2.77	Gur	533	0.64	21600
F10D10C8S12	M	0.18	3.53	Gur	537	1.51	4689
F10D10C8S30	I	0.43	3.33	Gur	1118	2.65	1672
F10D10C8S30	M	0.75	4.26	Gur	1084	2.21	4605
F10D10C8S90	I	-	2.48	Gur	5658	2.40	2254
F10D10C8S90	M	-	4.76	Gur	5483	4.49	1046
F10D10C8S120	I	-	4.28	LR	4715	1.66	21600
F10D10C8S120	M	-	4.28	LR	7087	4.28	21600
F10D10C8S300	I	-	3.65	LR	12919	-	21600
F10D10C8S300	M	-	3.31	LR	7521	-	21600
F10D20C8S6	I	0.25	2.77	Gur	487	1.28	1497
F10D20C8S6	M	0.08	1.23	Gur	494	0.66	306
F10D20C8S12	I	0.5	3.74	Gur	756	1.61	2496
F10D20C8S12	M	0.39	2.31	Gur	710	0.87	10761
F10D20C8S30	I	0.9	4.04	Gur	2568	3.75	870
F10D20C8S30	M	-	2.18	Gur	2621	1.56	7649
F10D20C8S90	I	-	3.67	Gur	8013	2.85	21600
F10D20C8S90	M	-	2.70	Gur	8326	2.07	21600
F10D20C8S120	I	-	3.90	LR	11439	3.77	21600
F10D20C8S120	M	-	2.66	LR	11808	2.66	21600
F10D20C8S300	I	-	4.03	LR	21600	-	21600
F10D20C8S300	M	-	2.96	LR	21600	-	21600
F17D50C8S6	I	0.46	2.39	Gur	225	0.92	5838
F17D50C8S6	M	0.28	3.03	Gur	198	0.74	21600
F17D50C8S12	I	-	1.83	Gur	1539	1.24	21600
F17D50C8S12	M	-	2.25	Gur	1186	1.74	2426
F17D50C8S30	I	-	2.36	LR	3316	1.40	21600
F17D50C8S30	M	-	2.28	LR	3458	-	21600
F17D50C8S90	I	-	2.50	LR	9367	-	21600
F17D50C8S90	M	-	3.31	LR	10604	-	21600
F17D50C8S120	I	-	2.87	LR	12817	-	21600
F17D50C8S120	M	-	3.23	LR	13453	-	21600
F17D50C8S300	I	-	2.91	LR	21600	-	21600
F17D50C8S300	M	-	3.60	LR	21600	-	21600
F17D70C8S6	I	-	1.25	Gur	825	0.68	5475
F17D70C8S6	M	3.98	4.36	Gur	249	0.59	21600
F17D70C8S12	I	-	4.46	LR	503	1.96	21600
F17D70C8S12	M	-	2.33	Gur	517	0.77	21600
F17D70C8S30	I	-	2.03	LR	4295	0.83	4812
F17D70C8S30	M	-	2.49	LR	3923	-	21600
F17D70C8S90	I	-	3.33	LR	15140	-	21600
F17D70C8S90	M	-	2.69	LR	11249	-	21600
F17D70C8S120	I	-	2.98	LR	18108	-	21600
F17D70C8S120	M	-	2.81	LR	14191	-	21600
F17D70C8S300	I	-	3.36	LR	21600	-	21600
F17D70C8S300	M	-	3.46	LR	21600	-	21600

Table 2: Performance comparison of Gurobi, the Lagrangian heuristics, and the Lagrangian heuristic with subsequent restricted MIP model. An “-” for the restricted MIP indicates that it did not improve over the Lagrangian heuristic solution.

Instance	Adj. costs	Dist. limit	Gur 24h Gap [%]	LR 6h Gap [%]	Best bound	Converge Time [s]	R-MIP Gap [%]	R-MIP Time [s]
F10D10C8S12	1	1	0.18	3.53	Gur	537	1.51	4689
F10D10C8S12	2	1	0.45	3.05	Gur	561	0.85	21600
F10D10C8S12	5	1	0.26	3.46	Gur	557	0.77	21600
F10D10C8S12	10	1	0.26	3.74	Gur	532	0.72	21600
F17D10C8S12	1	1	0.25	4.45	Gur	513	1.02	21600
F17D10C8S12	2	1	1.10	2.61	Gur	662	1.55	21600
F17D10C8S12	5	1	1.08	3.82	Gur	650	1.59	21600
F17D10C8S12	10	1	1.03	4.27	Gur	745	1.62	21600
F17D10C8S12	1	0.5	0.34	3.79	Gur	561	1.07	21600
F17D10C8S12	2	0.5	1.64	2.91	Gur	702	2.26	21600
F17D10C8S12	5	0.5	1.62	3.71	Gur	560	2.29	21600
F17D10C8S12	10	0.5	1.55	4.41	Gur	630	2.37	21600
F17D70C16S120	1	1	-	1.50	LR	21600	-	21600
F17D70C20S120	1	1	-	2.02	LR	21600	-	21600
F17D70C16S120*	1	1	-	2.90	LR	21600	-	21600
F17D70C20S120*	1	1	-	3.52	LR	21600	-	21600

Table 3: Computational results

7 Conclusions

We have presented a general formulation of a multi-stage stochastic facility location problem with capacity adjustments. The RVMSS indicates the economical benefit of the multi-stage stochastic formulation over the rolling horizon deterministic formulation if capacity adjustments become more expensive. Limiting the distribution distance also further increases the RVMSS.

Given the difficulty of solving the multi-stage planning problem, even with state-of-the-art general-purpose solvers, we have further presented a solution method based on Lagrangian relaxation. While Gurobi can solve only very small problems with a low number of scenarios, the proposed Lagrangian relaxation method is capable of solving real-world sized problems with a significantly higher number of scenarios, generally providing high quality solutions within 6 hours. For smaller instances, a subsequent restricted MIP approach further improves the solution quality. To solve the Lagrangian dual, we present a hybrid approach, switching from a cutting plane method with box constraints to the classical subgradient method to update the Lagrangian multipliers after a fixed number of iterations. This enables the method to decrease the computational time significantly, since solving the cutting planes method in later iterations becomes more time-consuming during the iterative solution process.

In future work, different heuristics to improve solutions for larger instances can be studied. Further, scenario reduction techniques can be applied to decrease the problem size.

References

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.

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

Allman, A. and Zhang, Q. (2020). Dynamic location of modular manufacturing facilities with relocation of individual modules. *European Journal of Operational Research*, 286(2):494–507.

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

Antunes, A. and Peeters, D. (2001). On solving complex multi-period location models using simulated annealing. *European Journal of Operational Research*, 130(1):190–201.

Bakker, H., Dunke, F., and Nickel, S. (2020). A structuring review on multi-stage optimization under uncertainty: Aligning concepts from theory and practice. *Omega*, 96:102080.

Becker, T., Lier, S., and Werners, B. (2019). Value of modular production concepts in future chemical industry production networks. *European Journal of Operational Research*, 276(3):957–970.

Birge, J. R. and Louveaux, F. (2011). *Introduction to Stochastic Programming.* Springer Series in Operations Research and Financial Engineering. Springer, New York, NY, 2 edition.

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., Melo, T., and Saldanha-da Gama, F. (2013). Comparing classical performance measures for a multi-period, two-echelon supply chain network design problem with sizing decisions. *Computers & Industrial Engineering*, 64(1):366–380.

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.

Cortinhal, M. J., Lopes, M. J., and Melo, M. T. (2015). Dynamic design and re-design of multi-echelon, multi-product logistics networks with outsourcing opportunities: A computational study. *Computers & Industrial Engineering*, 90:118–131.

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).

Escudero, L. F., Garín, A., Merino, M., and Pérez, G. (2007). The value of the stochastic solution in multistageproblems. *TOP*, 15(1):48–64.

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.

Huang, K. and Ahmed, S. (2009). The value of multistage stochastic programming in capacity planning under uncertainty. *Operations Research*, 57(4):893–904.

Huangfu, Q. and Hall, J. A. J. (2018). Parallelizing the dual revised simplex method. *Mathematical Programming Computation*, 10:48–64.

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.

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 Da-Gama, F. S. (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.

Ministry of Petroleum and Energy (2023). Norway's energy supply system: Electricity production. `https://energifaktanorge.no/en/norsk-energiforsyning/kraftproduksjon/`. last accessed 13.11.2023.

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.

Nickel, S., Saldanha-da Gama, F., and Ziegler, H.-P. (2012). A multi-stage stochastic supply network design problem with financial decisions and risk management. *Omega*, 40(5):511–524.

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

Ortiz-Astorquiza, C., Contreras, I., and Laporte, G. (2018). Multi-level facility location problems. *European Journal of Operational Research*, 267(3):791–805.

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.

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. and Tomasgard, A. (2011). The impact of flexibility on operational supply planning. *International Journal of Production Economics*, 134(2):300–311.

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.

Singh, K. J., Philpott, A. B., and Wood, R. K. (2009). Dantzig-wolfe decomposition for solving multistage stochastic capacity-planning problems. *Operations Research*, 57(5):1271–1286.

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. (2022). Locating hydrogen production in norway under uncertainty. In Ramalhinho, H., De Armas, J., and Voß, S., editors, *Computational Logistics*, volume 13557, pages 306–321, Cham. Springer.

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, Cham. Springer.

Štádlerová, Š., Schütz, P., and Tomasgard, A. (2024). Multi-period facility location and capacity expansion with modular capacities and convex short-term costs. *Computers & Operations Research*, 163:106395.

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.

Williams, H. P. (2013). *Model building in mathematical programming*. John Wiley & Sons, 5 edition.

A Hydrogen demand

Figure 5 shows the maximum aggregated daily demand for each of the demand components over the planning horizon. The deterministic component represents the demand from maritime passenger transportation in Norway (Danebergs and Aarskog, 2020). This demand is aggregated over 51 customer locations in Norwegian ports. Component "Stochastic 1" represents the demand from land-based transportation and is based on DNV GL (2019). Here we consider a total of 70 customer locations: 51 demand points located in Norwegian ports plus additional 19 municipalities with the highest road traffic volumes according to Statistics Norway (2018). Demand from land-based transportation is then divided among the 70 customers according to the relative traffic volume. Finally, component "Stochastic 2" represents demand from offshore operations (Ocean Hyway Cluster, 2020; Aglen and Hofstad, 2022). This demand is distributed among the 51 maritime customer locations.

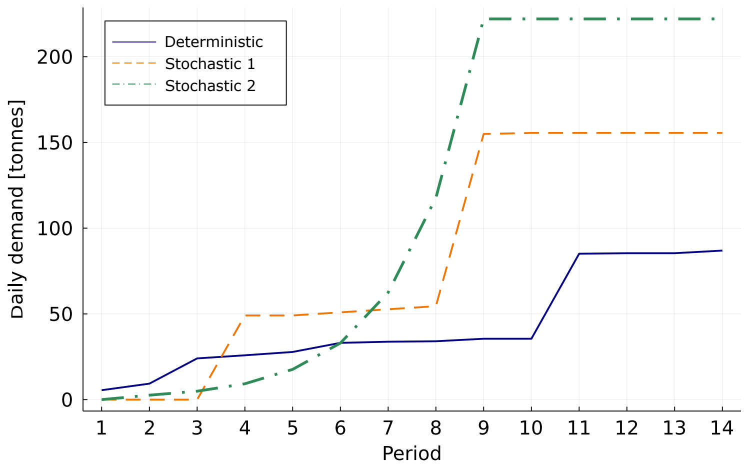


Figure 5: Maximum daily demand of each demand component (Štádlerová et al., 2023)

B Case study cost data

Table 4 shows investment and production costs for each capacity level. The production costs are considered at 100% capacity utilization, whereas "Production S" and "Production N" refer to production costs in the southern part of Norway and the northern part of Norway, respectively.

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
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 4: Investment and production costs at 100% utilization (Štádlerová et al., 2023)

Table 5 shows hydrogen distribution costs per kilometre and kilogram of hydrogen, whereas the unit price depends on the distance travelled.

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: Distribution costs in [€/km/kg H_2] (Štádlerová and Schütz, 2021)