The ATHENA framework: Analysis and design of a strategic hydrogen refuelling infrastructure

Citation for published version:

Digital Object Identifier (DOI):
10.1016/j.ijhydene.2023.04.190

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Publisher's PDF, also known as Version of record

Published In:
International Journal of Hydrogen Energy

Publisher Rights Statement:
© 2023 The Author(s).

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The ATHENA framework: Analysis and design of a strategic hydrogen refuelling infrastructure

Christa Searle, Kit Daniel Searle, Senlei Wang, Philip Greening

Centre for Logistics and Sustainability, Edinburgh Business School, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom
School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom

**Highlights**
- Novel framework to design a robust hydrogen refuelling infrastructure.
- Spatio-temporal mapping of hydrogen demand using logistics flow density.
- MILP optimisation for station locations, capacities, hydrogen supply and delivery.
- Agent-based simulation model for analysis with stochastic hydrogen demand.
- Case study for rollout and mature phase for heavy good vehicles in Northern England.

**Abstract**
With the pressured timescale in determining effective and viable net zero solutions within the transport sector, it is important to understand the extent of implementing a new refuelling infrastructure for alternative fuel, such as hydrogen. The proposed ATHENA framework entails three components which encapsulates the demand data analysis, an optimisation model in determining the minimal cost hydrogen refuelling infrastructure design, and an agent-based model simulating the operational system. As a case study, the ATHENA framework is applied to Northern England focusing on the design of a hydrogen refuelling infrastructure for heavy goods vehicles. Analysis is performed in calibrating parameters and investigating different scenarios within the optimisation and agent-based simulation models. For this case study, the system optimality is limited by the feasible number of tube trailer deliveries per day which suggests an opportunity for alternative delivery methods.

© 2023 The Author(s). Published by Elsevier Ltd on behalf of Hydrogen Energy Publications LLC. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

**Introduction**
There is a push within different sectors and across countries to reach climate change targets. The transportation sector is fundamental to human society, however, it is currently one of the largest global carbon dioxide emitting sectors [27]. Within this sector, the advances of innovative fuel solutions, such as hydrogen, biofuels and batteries, will play a significant role in the reduction of carbon dioxide emissions. The adoption towards such a strategy, however, is often seen as a challenge in...
having to consider both the vehicle commercialisation as well as the refuelling infrastructure which may be required to serve the former.

When considering the case for an alternative fuel, such as hydrogen, primary questions that may arise in the design of the refuelling infrastructure include “What are the refuelling demands?” “What is the minimum number of refuelling stations required?” and “Where should the hydrogen refuelling stations be geographically located?” Furthermore, it is critical to understand the investment required for the hydrogen refuelling and distribution networks. Hydrogen cost competitiveness typically depends on various aspects, such as the ability to scale in size and application, as well as the usage of existing pipelines, although it is only considered to reach competitiveness by 2030 [11].

Green et al. [9] highlight the challenges for the design and deployment of hydrogen refuelling infrastructures and allude to the existing methodologies in determining the locations for refuelling stations. A spatial geographic information systems approach to this was applied by Brey et al. [5]; while Sun et al. [29] combine the spatial factors in a mathematical model using particle swarm optimisation. In a study by Thiel [31], the particle swarm optimisation technique is further combined with agent-based simulation modelling in solving a price-based location model. Kuvvetli [14] and Kim et al. [13] follow a multi-objective approach in determining the locations of stations with case studies in Turkey and Korea, respectively. Similar to Kuvvetli, Gec¸ici et al. [8] consider a multi-period approach to consider the deployment of the infrastructure over time with a case study in Turkey. Other case-specific techno-economic analyses for the deployment of a hydrogen refuelling infrastructure include the studies by Ayodele et al. [3] and Simunović et al. [37] as applied to South Africa and Croatia, respectively.

Although research has been conducted in locating hydrogen refuelling stations (refer to Lin et al. [18] for a further review of hydrogen refuelling station location models), research on the complete design of a hydrogen refuelling infrastructure considering localised and centralised hydrogen supply configurations with different hydrogen transportation and delivery modes is less common and examples in the literature are case specific. Therefore, there exists a gap in the literature for a framework in conducting the analysis and design of a hydrogen refuelling network on strategic and operational levels.

The key contributions of this paper are as follows: A novel framework encapsulating three key phases in delivering a robust hydrogen refuelling infrastructure network design is proposed. This framework is then rendered useful by means of a case study whereby a novel methodology is proposed for creating a spatio-temporal mapping of hydrogen refuelling demand using logistics flow density, where granular refuelling data are not available. Within the case study, an optimisation model is employed to determine the hydrogen refuelling infrastructure for heavy goods vehicles in Northern England with the location, capacity and hydrogen supply configurations for each hydrogen refuelling station. Finally, a novel agent-based simulation model is developed for the case study which allows for the incorporation of stochasticity and simulating of operational decisions within the hydrogen refuelling infrastructure.

In addition to the introductory section, the remainder of this paper is structured as follows: In the next section, the methodology, i.e. the proposed ATHENA framework, is described in detail and in section thereafter, the framework is applied to a case study for heavy goods vehicles in Northern England and the results and limitations are discussed. This is followed by the concluding discussion in the final section.

Methodology: the proposed ATHENA framework

In light of the aforementioned gap in the literature, we propose a generic framework for the analysis and design of a strategic hydrogen refuelling infrastructure, called ATHENA. The framework, as illustrated in Fig. 1, entails three main components — the demand data analysis, an optimisation model, and an agent-based simulation model. The first two components form part of the strategic design, whereas the final component allows for operational design.

The input data required for each of the components are illustrated with dotted lines. The output of the first component (demand data analysis) is a spatio-temporal mapping of the hydrogen demand, which feeds into the second component (optimisation model). The output from this component is a hydrogen refuelling infrastructure network. This network may further be analysed in an operational manner in the third component (agent-based simulation model) to ensure a robust refuelling infrastructure. The framework may be applied to a specific region or geographic area under consideration and may involve one or more transportation modes. The remainder of this section is dedicated to provide a detailed description of each of the components.

Strategic demand data analysis component

The demand data analysis component employs input related to the logistics movement (such as road density flow rates, an origin-destination matrix of demand) and the overall refuelling demand of the transportation mode under consideration, together along with a set of existing and potential refuelling sites for each mode, together called the set of candidate hydrogen refuelling sites, and their geographical locations. In an ideal situation, the hydrogen refuelling demand at each candidate refuelling site would be known, however, such data are often not available and need to be estimated based on logistics movement, trends of existing fuel demand and adoption rates to hydrogen as fuel. Moreover, it is worth noting that space and time are inseparable and therefore the design of a hydrogen refuelling infrastructure (or any other refuelling infrastructure, for that matter) is inherently a spatio-temporal process. Therefore, the primary objective of the demand data analysis component is to create a spatio-temporal mapping of hydrogen refuelling demand for the transportation modes under investigation based on the aforementioned input data.

In light of this, we consider two distinct mappings, a temporal mapping and a spatial mapping, as graphically illustrated in Fig. 2. In general, we would expect the temporal variation in hydrogen refuelling demand to be brought about by the market adoption and technology readiness. A challenge
that exist is the co-evolution of the hydrogen refuelling infrastructure and the adoption of hydrogen fuel cell vehicle [9]. In the early stages of the establishment of a hydrogen refuelling infrastructure, a number of refuelling stations may be rolled out at critical locations to initiate the market growth. This growth may further be supported by government policies and regulations (in reaching climate targets), large-scale investments, and key entities which may spur action in the public and private sectors [12]. Hydrogen also competes with other alternative fuel solutions, such as battery electric

Fig. 1 – The proposed ATHENA framework.

Fig. 2 – An illustration of the temporal mapping and spatial mapping to depict the hydrogen refuelling demand.
vehicles and therefore, the adoption to hydrogen may be directly effected by the adoption and technology readiness of these alternatives.

The temporal mapping is derived from the total existing refuelling demand to a hydrogen demand temporally based on the conversion and the adoption rates of hydrogen refuelling vehicles. These rates may differ significantly across different transportation modes. As such, the overall forecasted refuelling demand of the existing fuel may be considered for each mode. In turn, this existing refuelling demand should be converted to a hydrogen fuel demand, taking into consideration the fuel efficiency of the existing fuel and fuel cell vehicles. For each time instance considered, the associated adoption rate for conversion to hydrogen for each transportation mode should be estimated. With these adoption rates, the hydrogen refuelling demand per mode may be estimated for each time instance to create a temporal mapping as illustrated graphically in Fig. 2.

Spatial variations in the refuelling demand may be brought about the traffic density of surrounding major roads, the population density of the surrounding towns or cities, the proximity to major industrial areas and government incentives, to name a few. The spatial mapping is the mapping of the estimated hydrogen demand per mode to the set of candidate sites based on geographical considerations, for each time instance. The temporal mapping therefore serves as input to the spatial mapping process by providing the estimated hydrogen demand per mode for each time instance. Additional input may include data on the logistics movement, such as flow data or origin-destination matrices, demographic and geographic information and insight from experts, such as to understand the key geographical considerations for hydrogen adoption. Other considerations may include the possibility of certain candidate sites being able to serve more than one mode.

Data analytical techniques and machine learning may be employed in an attempt to determine the refuelling demand at each of the candidate refuelling sites. This may then be performed for each of the time instances. Depending on the time horizon of the temporal mapping, the spatio-temporal mapping can be classified as either an aggregated mapping or a disaggregated mapping. In the former case, the mapping from one time period to another will depend on the mappings in the previous time instances. This would typically be the case if the time instances are in short intervals (typically a time interval of a few years). In the latter case, the mapping from one time period to another will not depend on the mappings in the previous time instances. This would typically be the case if the time instances are in long intervals (typically a time interval of decades or more).

The output of the demand data analysis entails a spatio-temporal mapping of the hydrogen refuelling demand across the set of candidate sites within the specified area for each transportation mode under consideration. This spatio-temporal analysis of the hydrogen demand may provide visual insight as to the spread of the geographic demand over time again illustrated graphically in Fig. 2. Note that the map in the figure is representative of any geographic area that may be considered.

**Strategic optimisation model component**

The second component of the ATHENA framework entails an optimisation model for the strategic design of a hydrogen refuelling infrastructure. The output of the spatio-temporal analysis performed in the first component, i.e. a set of candidate refuelling sites each with an associated time-dependent demand, serves as input to the optimisation model, along with input related to the production supply of hydrogen, the distribution of hydrogen and all of the associated capital and operational costs. The output of the optimisation model is a set of active hydrogen refuelling stations, each with an associated storage capacity and hydrogen supply method, which we call the optimal hydrogen refuelling infrastructure.

In general, this problem may be classified as a discrete facility location problem [1,17,21]. According to Laporte et al. [15]; a facility location problem comprises of two primary decisions, namely location and allocation decisions. The location decisions are concerned with determining the location of a facility, in this case the locations of hydrogen refuelling stations, while the allocation decisions are associated with determining how each facility should be supplied, in this case how the hydrogen should be supplied to the hydrogen refuelling stations. In the remainder of this paper, we refer to the allocation decisions as hydrogen supply decisions. Furthermore, facility location problems may be classified as multi-period facility location problems if there is a temporal component. In the remainder of this section, we discuss the location, allocation and temporal considerations together with the considerations for the objective function in more detail.

**Location considerations**

The primary objective of the location considerations is to determine which of the set of candidate refuelling sites should be selected as active refuelling stations. Perhaps the most important consideration is whether the demand at each candidate refuelling site must be satisfied and, more generally, if all of the hydrogen demand must be satisfied.

In the case where not all of the hydrogen demand must be satisfied, it is important to determine whether the number of hydrogen refuelling stations is pre-determined. If this is the case, then the p-median facility location problem would be a natural point of departure where p is the number of hydrogen refuelling stations to be activated. If the p-median facility location problem is employed then there may be two possible objectives, to minimise the unsatisfied demand or to minimise the cost of satisfying the demand. If the number of hydrogen refuelling stations is not pre-determined, then the fixed charge facility location problem would be a natural point of departure. In this case, there is a fixed cost associated with activating a candidate hydrogen refuelling site. This fixed cost may be constant or it can be broken down into a standard fixed cost for establishing the infrastructure together with a variable cost depending on the size of the hydrogen refuelling station. The objective function of the fixed-charge facility location problem is typically to minimise the cost of establishing the set of active refuelling stations together with the cost of supplying the hydrogen to the activated stations. If the fixed-charge facility allocation is employed then a constraint
specifying a lower-bound for the unsatisfied demand is required, otherwise the solution will be trivial.

If we consider the case where all of the hydrogen demand must be satisfied, then the fixed-charge facility location problem becomes trivial because the optimal solution would simply be to activate the entire set of candidate refuelling sites. In this case, however, we can employ a set cover problem to model the situation where the demand of one candidate hydrogen refuelling site may be satisfied at some neighbouring candidate hydrogen refuelling site provided that the distance between them is sufficiently small. In this case, if there is again a fixed and variable cost associated with activating a candidate hydrogen refuelling site, the decisions then become which candidate hydrogen refuelling sites to activate and which inactive sites each must cover. The objective function in this situation is typically to minimise the cost of establishing the hydrogen infrastructure and supplying the hydrogen to the active refuelling stations. Within the set cover facility location framework, it is important to include an upper-bound on the capacity of an active hydrogen refuelling station to ensure that a single active refuelling station does not become too large.

Hydrogen supply considerations

One of the first arguments typically made when considering hydrogen production and supply is which type of hydrogen production, such as blue or green, to employ. In this paper we do not consider this argument as, for the purpose of the optimisation model, the different production types of hydrogen are indistinguishable (the type of hydrogen will only change the user-determined parameters associated with production and set-up costs). In this section we are more concerned with how the produced hydrogen should be supplied and delivered to the active hydrogen refuelling stations.

In general there are two schools of thought when considering the supply of hydrogen, namely localised production and centralised production with transportation [2]. In the former case, the hydrogen is produced locally at the hydrogen refuelling station using production technologies, such as steam methane reforming or water electrolysis. Producing hydrogen locally on-site has the advantage that the hydrogen produced can directly be dispensed to potential customers without having to be transported to the station. If an on-site production technology is employed then there are typically fixed costs associated with the equipment for producing, purifying and managing the hydrogen prior to storage (also known as capital expenditure) together with a variable cost associated with producing a unit of hydrogen (also known as operational expenditure) [2]. The capital costs of the on-site production technologies will typically depend on the production capacity and installation costs of the associated equipment, while the operational expenditure typically depends on the general operations, maintenance costs and raw material costs [6]. Therefore, for localised on-site hydrogen production a set of different technology types, each with an associated capital cost, operational cost and production capacity, may need to be specified.

Producing hydrogen at a centralised production facility, on the other hand, has the advantage of economies of scale. By this we imply that, given a large enough demand, by having a few large centralised production facilities we may be able to produce all of the required hydrogen at a lower operational cost and potentially a lower capital expenditure cost [9]. In a similar manner to the on-site localised production, there is generally a fixed cost associated with opening a centralised production plant, which again may depend on the production capacity of the facility, together with a variable cost associated with producing a unit of hydrogen. Therefore, for centralised off-site production, a set of candidate centralised production facilities, each with an associated capital cost, operational cost and production capacity may need to be specified.

Producing hydrogen at a centralised off-site production facility has the drawback that there is an associated cost with transporting the hydrogen to the required refuelling station. The primary methods for the distribution of hydrogen from centralised production facilities or storage facilities to the hydrogen refuelling stations include tankers (for liquid hydrogen), tube trailers and a hydrogen pipeline (the latter two for hydrogen in gaseous form). In the case of the tankers and tube trailers, the capacity per vehicle and number of trips that may be performed per day are constraints that should be considered in the model. Moreover, the cost of transporting the hydrogen by means of tanker or tube trailer may depend on the capacity of the vehicle and the distance that the vehicle must travel. In some cases there may be a capital cost associated with purchasing a fleet of delivery vehicles, but the capital costs are often included in the operational costs. For the hydrogen pipeline, there may be considerations towards the pipeline involving any blending with another gas, the purification that may be required, and the location of the pipeline. There may be an existing or proposed pipeline which may be added to the optimisation model or, alternatively, the model may be expanded to determine the location of the hydrogen pipeline.

Finally, when deriving the supply conditions it may be specified whether only a single or both production and supply methods may be employed to serve the demand at a single active refuelling station.

Objective function considerations

Perhaps the most important component of any optimisation model is the objective function. In each of the previous sections we have discussed the modelling considerations with a view of an objective function which minimises the overall infrastructure and hydrogen supply cost. At this point, we note that the objective function may take on various other forms. The most interesting of these may be to determine the environmental impact of establishing a hydrogen refuelling infrastructure which minimises the environmental impact of establishing a hydrogen refuelling infrastructure [23,26]. For this case, instead of supplying a cost associated to each of the aforementioned design considerations we would be required to supply the environmental impact associated with each decision. There may, however, still be a cost associated with all of the infrastructure set-up and hydrogen supply cost which may be included in the form of a budget constraint. Another interesting objective function may be to minimise the safety risk associated with establishing the hydrogen refuelling infrastructure. At this point, we also note that all three of the aforementioned
decisions can be considered together if a multi-objective optimisation approach is adopted.

Temporal considerations
In the previous three subsections within the ‘Strategic optimisation model component’, we have limited our conversation on the location, hydrogen supply and objective function considerations to a single time period. The input to the strategic optimisation model is, however, a spatio-temporal mapping of hydrogen refuelling demand and therefore, it is important to consider how the temporal elements may be incorporated in the model. Therefore, in this section we discuss the temporal considerations from the viewpoint of a multi-period facility location problem [20]. In a very similar fashion to the demand data analysis component, there are two general strategies for handling multi-period facility location problems, an aggregated approach and a disaggregated approach.

The disaggregated approach is perhaps seen as the simplest because the decisions in one time instance do not affect the decisions in a later time instance. In this case, the selected model can be formulated as a single time-period facility location problem which can be solved for every time instance of the temporal mapping. This methodology is typically only adopted when the time instances are far apart as we might expect that decisions made during one time instance would have no effect on the later time instances.

On the other hand, in the aggregated approach the decisions made in one time instance may affect the decisions in later time instances. We discuss the aggregated temporal approach with respect to both the location and hydrogen supply considerations.

For the location decisions we may require, for example, that if a candidate hydrogen refuelling station is activated in a given time instance then it must remain active for the next, and possibly every, future time instance. In this case, if the number of active refuelling stations is pre-determined, then the multi-period p-median facility location approach may serve as a point of departure. In this case, it is important to ensure that the number of pre-determined active hydrogen refuelling stations does not decrease in each successive time instance. If the fixed charge facility location problem is adopted as the point of departure then the notion of an opening cost changes slightly with respect to the single time period fixed charge facility location problem. More specifically, for each candidate refuelling station, we may need to attach a fixed cost associated with installing the infrastructure required to open the hydrogen refuelling station together with another fixed cost associated with operating the hydrogen refuelling station in a given time instance. It may also be possible to include a separate decision that would allow us to close down an active hydrogen refuelling station, at an additional cost, in any given time instance. At this point, we also note that the set cover modelling methodology may be employed with the same aforementioned fixed costs, however, we may allow the coverage radius to change in each time instance.

For the hydrogen supply considerations we may need to consider if the hydrogen supply method may change from one time instance to another. Therefore, in a similar fashion to the location considerations, for each hydrogen supply type we may need to attach a fixed cost for establishing the technology together with a separate annual operational cost which may or may not depend on the quantity of hydrogen supplied. If we are allowed to change the hydrogen supply method, then we may need to include a cost for this change from one time instance to another. Depending on the temporal scale, we may need to consider the life cycle of the different hydrogen production technologies. This may specifically be important when considering localised production.

To summarise, the output of the strategic optimisation model component of the framework includes a validated optimisation model for a hydrogen refuelling infrastructure network which may include certain analysis, and the output of the model detailing the refuelling infrastructure specifications. This, together with the demand data analysis component, comprises the strategic design of the hydrogen refuelling infrastructure.

Operational agent-based simulation model component
The final component of the ATHENA framework entails an agent-based simulation model that relates to the operational design of the hydrogen refuelling infrastructure. The output from the second component of the ATHENA model, which is an optimal hydrogen refuelling infrastructure (i.e. a set of active hydrogen refuelling stations each with an associated demand and hydrogen supply method) serves as the input to the agent-based simulation model. Additional input includes operational factors of a hydrogen refuelling infrastructure, that may be of interest to model in answering operational design questions, such as the reorder policy or the maintenance schedule. The output of the agent-based simulation model component is a model with which experiments may be performed to ensure a robust operational hydrogen refuelling infrastructure.

Agent-based modelling allows for a bottom-up approach in which the decisions, behaviour and interactions of individual agents are simulated over time which may lead to an emergent large-scale system behaviour that may be captured [19]. The agent classes may include the set of active hydrogen refuelling stations, the production facilities, the delivery fleet transporting the hydrogen, and the customers or end users. Agent-based modelling further allows for uncertainty to be incorporated in the model and for different scenarios to be simulated in investigating and analysing the system. The agent-based simulation model may therefore act as validation of the hydrogen refuelling infrastructure output from the optimisation model component of the ATHENA framework, while further addressing specific operational questions.

An agent-based simulation model is able to imitate the operations within a typical refuelling infrastructure, such as the end users refuelling at stations, the transporting of hydrogen (e.g. using tube trailers or tankers) from centralised production facilities to hydrogen refuelling stations, and the operations at the refuelling stations while keeping track of the quantity of hydrogen available. The model may include a demand profile (either deterministic or probabilistic) for each end user or for each hydrogen refuelling station in any level of detail. In the case where the end user (i.e. a heavy goods vehicle driver) is an agent, decisions to consider may include when they need to refuel, how they choose where to refuel
and if they arrive at a station that is unable to serve them, which station will they choose to visit next.

In the case where each hydrogen refuelling station is an agent, the stock levels of the hydrogen available may be monitored and experiments may be performed as to the replenishment or reorder policy for each station. Other decisions pertaining to a refuelling station as an agent may involve the storage capacity, capacity of the on-site production technology (in the event of localised production), number of refuelling points and maintenance schedule. The centralised production facilities may also act as agents with certain attributes, parameters and policies. This may involve the production capacity, the storage availability, the number of tankers or tube trailers that may be served at a time, and the maintenance schedule. For thetube trailer or tanker as an agent, the attributes, parameters and decision making may involve the capacity, distance to travel (monitoring of fuel levels), and routing decisions.

The agent-based model may act as a decision support tool and provide an environment within which the impact of different operational uncertainties in the hydrogen refuelling infrastructure may be analysed. The model may allow for the response of the system to adjustments in certain key operational parameters to be captured and allow for parameters variation and sensitivity analysis in determining the configuration of the system for a cost-effective and robust hydrogen refuelling infrastructure. Generally, we aim to identify the configuration that results in the lowest cost for the hydrogen refuelling infrastructure and a higher refuelling efficiency in terms of the served refuelling demand at hydrogen refuelling stations. The use of reinforcement learning within an agent-based simulation model may further allow for a self-organising system to determine effective operational solutions for the model.

The output of operational the agent-based simulation model component of the ATHENA framework is therefore a validated agent-based model, as well as a robust hydrogen refuelling infrastructure supplementing to the output of the optimisation model.

### Case study

The ATHENA framework, described in the previous section ‘Methodology: The proposed ATHENA framework’, is applied to a real-life case study in Northern England. Given the above average carbon intensive economy in Northern England, compared to other regions in the country, the opportunity for hydrogen has been identified [16]. Serving as validation to the proposed ATHENA framework, this case study is a model concept demonstrator that is designed and developed to recommend a hydrogen refuelling infrastructure network to serve heavy goods vehicles in this region.

### Strategic demand data analysis

The first component of the ATHENA framework is the demand data analysis, which is conducted at the hand of a temporal mapping and spatial mapping, as described in the subsection ‘Strategic demand data analysis component’. Two distinct time instances are considered for the temporal mapping, that is the rollout phase as estimated in 2027, and the mature phase as estimated in 2040. We call $T$ the set of time instances.

The first step in the temporal mapping is to convert the existing refuelling demand in the Northern England region and translate this to an equivalent hydrogen demand. Since the majority of the refuelling demand for heavy goods vehicles is derived from diesel, we do not consider any alternative fuels in this case study. To this end, data were collected from publicly available sources and through conversations with regional transport bodies to estimate the total existing diesel fuel demand for heavy goods vehicles in Northern England [33,35]. As quantitative data are limited, this diesel refuelling demand is expected to be 41% of the total diesel heavy goods vehicle refuelling demand in Great Britain based on the number of trips that pass through the North of England [33]. More specifically, we estimate that the total annual diesel demand in the Northern England region is $2330408731$. Since it is very difficult to forecast the future diesel refuelling demand for various reasons, we assume that this value remains unchanged for each of the time instances. Next, we consider the temporal mapping of this diesel refuelling demand to a hydrogen refuelling demand for each time instance. First we translate the existing diesel refuelling demand in litres to the equivalent hydrogen refuelling demand in kilograms. To this end, if $D_j$ denotes the total refuelling demand for diesel (in litres) then the equivalent total hydrogen refuelling demand (in kilograms), denoted by $H_{j}$, can be expressed as

$$ H_j = (10 \psi_2)\left(\frac{\psi_0}{\psi_1}\right)^{33.33}. $$

(1)

where $\psi_0$ and $\psi_1$ denote the vehicle efficiency of a diesel and hydrogen heavy goods vehicle, respectively. The first bracket in (1) is employed to translate the total diesel demand in litres to a total energy demand in kWh, where we assume that 1 L of diesel amounts to 10 kW h. The second bracket in (1) is employed to translate the total energy requirement from diesel heavy goods vehicles to the required quantity of hydrogen in kilograms where we assume that 1 kg of hydrogen is equivalent to 33.33 kWh. In this case study we assume that $\psi_0 = 0.42$ and that $\psi_1 = 0.45$ [4,10] which results in a total hydrogen demand of $H_t = 652579703$. Note that this would be the total hydrogen refuelling demand if all of the current diesel refuelling demand is converted to hydrogen.

Therefore, for the temporal mapping, we assume that temporal variations will be brought about by the uptake of hydrogen fuel cell heavy goods vehicles. We assume that the adoption rate and conversion to hydrogen is estimated as 0.1% during the first time instance, the rollout phase, and 50% during the second time instance, the mature phase. These assumptions are informed by subject matter experts, although we realise that the adoption rate is a complex measure to estimate. Therefore, if $d_t$ denotes the total hydrogen demand during time instance $t \in T$ then

$$ d_t = \alpha t H_t, $$

(2)

where $\alpha_t$ denotes the adoption rate in time instance $t \in T$. Note that for the temporal mapping we employ a disaggregated approach since the two time instances are generally far apart.
Next, we consider the spatial mapping. In order to perform the spatial mapping, data on the candidate refuelling sites and logistics movement were collected. The set of candidate sites, denoted by \( I \), includes existing and potential sites for hydrogen fuel cell heavy goods vehicles to refuel within Northern England and amounted to a set of 158 warehouse facilities, 43 service stations, and 12 ports, denoted by \( I_W, I_S \) and \( I_P \), respectively. \[ I = I_W \cup I_S \cup I_P. \] Data on the journeys undertaken by heavy goods vehicles in the region are not available and therefore, alternative logistic flow data are considered along with qualitative knowledge on the refuelling behaviour of heavy goods vehicles.

We start off by employing a dataset on the heavy goods vehicle flow and density data for corridors in the region, as provided by Transport for the North [32], to determine, what we call a popularity estimate. This estimate for each candidate refuelling site is determined based on the density of the heavy goods vehicles traffic flow within a certain radius from the site. Therefore, the popularity estimate inherently assumes that candidate sites which are nearby to busy road segments would be a more popular refuelling site. We also assume that each candidate refuelling station type (i.e. warehouse facilities, service stations and ports) will satisfy a certain proportion of the total hydrogen refuelling demand. Therefore let \( \gamma_W, \gamma_S \) and \( \gamma_P \) be the relative proportions of the total hydrogen refuelling demand to be satisfied at warehouse facilities, refuelling stations and ports, respectively such that \( \gamma_W + \gamma_S + \gamma_P = 1 \). Furthermore, we also assumed that there exists a correlation between the size of a warehouse facility and its refuelling demand, and that the refuelling at warehouse facilities may be shared between heavy goods vehicles from different owners.

Finally, if we let \( h_t \) denote the hydrogen demand at candidate refuelling site \( i \in I \) at time instance \( t \in T \) then

\[
h^t_i = \begin{cases} 
\lambda_W \sum_{i \in I_W} s_i - \left(1 - \lambda_W \right) \sum_{i \in I_W} f_i & \text{if } i \in I_W \\
\lambda_S \sum_{i \in I_S} s_i + \lambda_S \sum_{i \in I_P} s_i - \left(1 - \lambda_S \right) \sum_{i \in I_S} f_i & \text{if } i \in I_S \\
\lambda_P \sum_{i \in I_P} s_i + \left(1 - \lambda_P \right) \sum_{i \in I_P} f_i & \text{if } i \in I_P
\end{cases}
\]

(3)

where \( s_i \) denotes the size of candidate site \( i \in I \), \( f_i \) is the maximum flow density of all the road segments within a specified radius of candidate site \( i \in I \) and \( \lambda_W, \lambda_S, \lambda_P \in [0, 1] \) are the relative importance weightings of the size and popularity factors. In (3), the terms \( \gamma_W d_i, \gamma_S d_i \) and \( \gamma_P d_i \) essentially determine the total hydrogen refuelling demand to be satisfied at warehouse facilities, refuelling stations and ports, respectively, during each time instance. This total hydrogen demand at each refuelling station type is then distributed amongst the set of candidate refuelling sites of that type proportionally to the size and popularity factors.

At this point we note that refinement of the demand mapping may be required, in the sense that the refuelling demand at candidate sites, which fall below some minimum demand threshold, is aggregated to the closest candidate site with a larger refuelling demand. This allows for a more realistic representation of the demand.

For this case study, we assume \( \lambda_W = 0.5 \) giving size and popularity factors equal importance for warehouse facilities, while \( \lambda_S = \lambda_P = 0 \) as the size factor is disregarded for the service stations and ports as the variation in size per station type for service stations and ports is assumed fairly small. Qualitative data from an entity representing fuel retailers estimated that 75% of the current diesel refuelling of heavy goods vehicles in Great Britain takes place at warehouse facilities [22]. As a result of the various industrial clusters based in Northern England, we increase this value such that \( \gamma_W = 0.8 \). We further assume that \( \gamma_S = 0.12 \) and \( \gamma_P = 0.08 \).

The spatio-temporal mapping for the hydrogen refuelling demand for heavy goods vehicles within Northern England is illustrated in Fig. 3, where Fig. 3(a) refers to the first time instance depicting a rollout phase and Fig. 3(b) refers to the second time instance depicting a mature phase. Each dot represents a candidate site, where the size and colour of the dot corresponds with the size of the refuelling demand at the candidate site. The mapping further provides insight as to the geographical distribution of the candidate sites. For this case, clustering of sites is visible and correlates with the geographical location of industrial clusters or hubs.

The spatial mapping for the first time instance, the rollout phase, contains 20 candidate sites of which 85% is warehouse facilities, 10% is port locations and 5% is service stations. The maximum daily demand allocated to a candidate site for this time instance is 148 kg, whereas the minimum and average daily demand allocated are 53 kg and 80 kg, respectively. For the second time instance, the mature phase, spatial mapping contains a set of 200 candidate sites of which 79% is warehouse facilities, 18% is service stations and 3% is port locations. The distribution of the daily demand of hydrogen across the set of candidate sites in the mature phase has a maximum of 19251 kg, a minimum of 578 kg and an average of 4470 kg. This spatio-temporal mapping, which provides a set of candidate sites and the associated refuelling demand for each candidate site, is the output from the demand data analysis component of the ATHENA framework.

**Strategic optimisation model**

The second phase of the ATHENA framework is to determine which of the set of candidate refuelling sites to establish as active hydrogen refuelling stations, and how hydrogen should be supplied to each active refuelling station. In the remainder of the discussion of this model, we essentially employ the model of Searle et al. [25].

For this case study, we employ a disaggregated approach and solve the model for each time instance. The model is formulated as a mixed integer linear programming model adopting a set-covering approach whereby the demand of a candidate refuelling site can be satisfied at some other candidate refuelling site provided that the distance between the two sites is less than the so-called maximum coverage distance. The decision variables of this model therefore include identifying which of the candidate refuelling sites are selected to be activated as hydrogen refuelling stations, the capacities of these hydrogen refuelling stations (based on the
hydrogen refuelling demand that is covered by the station) and then, for each hydrogen refuelling stations, whether hydrogen should be supplied from localised on-site and/or centralised off-site production, as well as the quantities of hydrogen to be supplied by either option. Additionally, if localised on-site production is selected at a hydrogen refuelling station, a decision variable to determine the size of electrolyser (we assume the use of water electrolysis) to be employed and, if centralised off-site production is selected, a decision variable is employed to determine the number of deliveries required along with the specific centralised production facility from which the hydrogen will be supplied.

The objective of the optimisation model is to minimise the overall cost of establishing the hydrogen refuelling stations and supplying the hydrogen. The objective function includes three primary elements relating to the hydrogen refuelling stations, as well as the centralised production and localised on-site production supply of hydrogen. The costs relating to establishing a hydrogen refuelling station includes a fixed setup cost and unit cost depending on the capacity of the hydrogen refuelling station. The cost of supplying hydrogen by means of a centralised production supply of hydrogen consists of the transportation cost and unit production cost is considered, while the cost for establishing localised on-site production supply of hydrogen consists of the fixed setup and unit production cost. At this point, we again note that a multi-objective approach can be adopted by extending the model of Searle et al. [25]. For the purpose of this case study, however, we consider the overall cost of the network to be the most important factor and therefore we employ the model as is.

For the application of the optimisation model to the Northern England region, data on the centralised production facilities are collected, along with data relating to the cost factors and the capacities of both localised on-site and centralised production as reported in Searle et al. [24]. The assumptions are derived from the UK’s Hydrogen Production Costs 2021 report [34] and informed by subject matter experts. Five centralised production facilities are identified primarily within the Teesside, Hornsea, Humber, Leeds, Merseyside areas. For both the rollout and mature phases, green hydrogen production is assumed at these facilities with no limitation to the production capacity. In the case of localised on-site production, the options of 1 MW, 5 MW, 10 MW, 30 MW and 50 MW electrolyisers are considered. The parameters that we employed for our implementation of the model are summarised in Tables 1–3.

For this case study, the optimisation model is validated through parameter variation and scenario analysis. A parameter variation is performed in which the maximum coverage distance parameter for both time instances, i.e. the rollout and mature phases, are varied. The output for the parameter variation is summarised in Fig. 4 where Fig. 4(a) refers to the first time instance depicting the rollout phase and Fig. 4(b) refers to the second time instance depicting the mature phase. In the rollout phase the maximum coverage distance is varied from 5 to 40 km, while in the mature phase this parameter is varied from 1 to 5 km. In both cases, the number of hydrogen refuelling stations decreases in a non-linear fashion as the maximum coverage distance increases. The maximum coverage distance parameter is calibrated and, for the application of the optimisation model to the Northern England region, data on the centralised production facilities are collected, along with data relating to the cost factors and the capacities of both localised on-site and centralised production as reported in Searle et al. [24]. The assumptions are derived from the UK’s Hydrogen Production Costs 2021 report [34] and informed by subject matter experts. Five centralised production facilities are identified primarily within the Teesside, Hornsea, Humber, Leeds, Merseyside areas. For both the rollout and mature phases, green hydrogen production is assumed at these facilities with no limitation to the production capacity. In the case of localised on-site production, the options of 1 MW, 5 MW, 10 MW, 30 MW and 50 MW electrolyisers are considered. The parameters that we employed for our implementation of the model are summarised in Tables 1–3.

For this case study, the optimisation model is validated through parameter variation and scenario analysis. A parameter variation is performed in which the maximum coverage distance parameter for both time instances, i.e. the rollout and mature phases, are varied. The output for the parameter variation is summarised in Fig. 4 where Fig. 4(a) refers to the first time instance depicting the rollout phase and Fig. 4(b) refers to the second time instance depicting the mature phase. In the rollout phase the maximum coverage distance is varied from 5 to 40 km, while in the mature phase this parameter is varied from 1 to 5 km. In both cases, the number of hydrogen refuelling stations decreases in a non-linear fashion as the maximum coverage distance increases. The maximum coverage distance parameter is calibrated and,
after consultation, the base case in the rollout phase is set to 20 km with the base case in the mature phase is set to 2 km.

The optimisation model output for the two time instances considered, the rollout and mature phases, are geographically illustrated in Figs. 5 and 6. The red icons indicate inactive candidate refuelling sites that are not activated as hydrogen refuelling stations, while the green and purple icons indicate activate hydrogen refuelling stations with centralised and localised on-site production supply, respectively. The blue icons represent the centralised production facilities, and the orange lines indicate the transportation link between hydrogen refuelling stations and centralised production facilities. Note that the case where a hydrogen refuelling station requires both localised on-site and centralised production supply is illustrated where a purple icon is linked by an orange line to a centralised production facility.

In the rollout phase, as graphically illustrated in Fig. 5, there are eight hydrogen refuelling stations which all make use of centralised hydrogen production supply, while in the mature phase illustrated in Fig. 6, there are 189 hydrogen refuelling stations of which 1% uses on-site hydrogen production supply, 24% uses only centralised hydrogen production supply and the remaining 74% uses both on-site and centralised hydrogen production supply. This high percentage of combined supply is due to certain binding constraints in the optimisation model which is further investigated by means of a sensitivity analysis.

A scenario analysis is performed to explore the on-site and centralised production alternatives in the mature phase and compare the objective function value of these scenarios with the base case from the optimisation output illustrated in Fig. 6. In Scenario A, the model only allows localised on-site production supply by not allowing any tube trailer deliveries to hydrogen refuelling stations, while in Scenario B there is no limitation to the number of tube trailer deliveries to hydrogen refuelling stations. In Scenario A, the overall cost increased from the base case by 18.11%, whereas in Scenario B the overall cost decreased from the base case by 31.44%. In the latter, however, the average number of tube trailer deliveries per day per hydrogen refuelling station exceeded nine which is not necessarily viable. This indicates that the system optimality is limited by the feasible number of daily deliveries, which motivates the implementation of alternative delivery methods, such as pipelines.

As an additional analysis, the effect of incrementally increasing number of allowed per hydrogen refuelling station is investigated with consideration to the portion of hydrogen refuelling stations across the different supply options. The analysis is shown in Fig. 7 with the average number of daily deliveries, as indicated by the curve on the graph, illustrating a non-linear increase.

This suggests that in the case with no limit on the maximum number of deliveries, the majority portion of hydrogen refuelling stations will have centralised production supply, however, there will still be a portion with localised on-site production supply.

<table>
<thead>
<tr>
<th>Table 2 – A summary of the maximum and minimum production capacity for the set of localised production types.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
</tr>
<tr>
<td>1 MW</td>
</tr>
<tr>
<td>5 MW</td>
</tr>
<tr>
<td>10 MW</td>
</tr>
<tr>
<td>30 MW</td>
</tr>
<tr>
<td>50 MW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3 – The fixed cost associated with each of the localised on-site production type for the rollout and mature phases.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
</tr>
<tr>
<td>1 MW</td>
</tr>
<tr>
<td>5 MW</td>
</tr>
<tr>
<td>10 MW</td>
</tr>
<tr>
<td>30 MW</td>
</tr>
<tr>
<td>50 MW</td>
</tr>
</tbody>
</table>

Fig. 4 – The analysis for varying the maximum coverage distance parameter within the optimisation model for the case study in Northern England for the rollout phase estimated in 2027 in Figure (a) and for the mature phase estimated in 2040 in Figure (b).
**Agent-based model with analysis**

The third and final phase of the ATHENA framework entails an agent-based simulation model, as described in the subsection ‘Operational agent-based simulation model component’. The agent-based model allows for the operational aspects of the hydrogen refuelling infrastructure to be modelled and, where the optimisation model considered a deterministic demand, the agent-based model has the ability of capturing the stochastic nature of refuelling demand which reflects reality more accurately when modelling operational decisions.

For the case study in Northern England, an agent-based simulation model is developed with three primary agent classes — the activated set of hydrogen refuelling stations, as determined in the optimisation model with analysis component of the ATHENA framework, the set of centralised production facilities, and tube trailers delivering the hydrogen supplied centrally to refuelling stations.

The daily hydrogen production at each of the centralised production facilities is modelled and the level of hydrogen available is monitored as it increases with production and decreases when hydrogen supply is delivered to hydrogen refuelling stations. The set of tube trailers are modelled such that each centralised production facility has a dedicated fleet of tube trailers that transports the hydrogen from the facility to the hydrogen refuelling stations which are supplied from that facility (as determined by the optimisation model in the subsection ‘Strategic optimisation model’). The number of journeys per tube trailer (when filled with hydrogen, and when empty) is monitored and as a tube trailer typically delivers a full load to a hydrogen refuelling station, they simply return to the centralised production facility after having made the delivery.

The set of hydrogen refuelling stations is modelled to each have a stochastic profile of hydrogen refuelling demand (thereby simulating the real-world demand from heavy goods vehicles refuelling at the station throughout the day) and a certain storage capacity for hydrogen at the station. Furthermore, the production supply of hydrogen to the refuelling station is modelled as determined by the optimisation model in the subsection ‘Strategic optimisation model’, whether from the centralised facility and delivered by tube trailers, or produced locally on-site. In the case of the former, a reorder policy is modelled for the hydrogen refuelling station to simulate the reorder point and quantity required. Finally, the level of hydrogen available at the refuelling station is monitored throughout the day.

The size of the storage at a hydrogen refuelling station directly influences the setup cost and therefore it is necessary to determine the size of the storage required in relation to the daily refuelling demand at the hydrogen refuelling station. For
the case study as applied to Northern England, an analysis is performed within the agent-based model to investigate the effect of the storage size on the cost and hydrogen refuelling demand. The optimisation model (in the subsection ‘Strategic optimisation model’) provides an upper bound defined as the scenario where all of the refuelling demand is met. The storage capacity of all of the hydrogen refuelling stations is varied as a percentage of the daily hydrogen refuelling demand at the station from 100% to 50% in increments of key performance indicators that are evaluated for each variation in the storage capacity of the stations are the percentage savings in hydrogen refuelling station variable cost and the percentage unsatisfied demand. The output of this parameter variation is illustrated graphically in Fig. 8.
It is clear that as the storage capacity of the hydrogen refuelling stations decreases, there is an increase observed in both the hydrogen refuelling station cost savings and the percentage of unsatisfied demand. There clearly exists a trade-off between saving costs with a smaller storage capacity and not being able to serve all of the refuelling demand. Further investigation may involve individual analysis of the storage capacity per refuelling station to determine a robust storage capacity required for each hydrogen refuelling station.

**Limitations**

In this section we note some of the limitations pertaining to the practical aspects of the case study. As mentioned in the subsection 'Strategic demand data analysis', two distinct time instances were considered — the rollout phase estimated in 2027 and the mature phase estimated in 2040. For the modelling of this temporal scale, however, a disaggregate approach was adopted (as mentioned in the subsections 'Strategic demand data analysis' and 'Strategic optimisation model') because it is assumed that the time instances are far apart. The limitation of this approach, as discussed in the subsection 'Temporal considerations', is that the decisions in the one time instance are assumed not to affect the decisions in the next time instance.

It is further noted that the parameters employed for the case study involve many assumptions due to the lack of available data. Where available, data were collected from various sources and subject matter experts. This, however, remains a limitation of the case study. The estimation of the cost of hydrogen production, which may be an influential factor, is especially challenging for a temporal scale considering more than a decade into the future.

Finally, the case study is limited in its scope. It considers only a single transportation mode (i.e. heavy goods vehicles), while in reality with the rollout of hydrogen multi-modal refuelling stations may exist and therefore need to be considered as a whole. Another limitation in scope of the case study is the distribution of hydrogen produced at central production facilities where we assumed that only tube trailers are employed. It is noted in the scenario analysis in the subsection 'Strategic optimisation model' that the system optimality is limited by the number of tube trailer deliveries per day.

**Conclusion**

In this paper, we propose the ATHENA framework for the analysis and design of a strategic hydrogen refuelling infrastructure. The framework consists of three components of which the first two, namely the demand data analysis and the optimisation model, serve to address the design of the refuelling infrastructure and the final component, namely the agent-based simulation model, serves to address the operational design. The aim of the demand data analysis component is to create a spatio-temporal mapping of the hydrogen refuelling demand. This demand then serves as input to the optimisation model component which aims to determine the optimal hydrogen refuelling infrastructure specifying the set of hydrogen refuelling stations, each with an associated capacity and hydrogen supply method. Finally, the agent-based simulation model component allows for the design and development of an agent-based model, based on the output from optimisation model, with which experiments may be performed to ensure a robust operational hydrogen refuelling infrastructure.

In validating the ATHENA framework, it was applied to a case study in Northern England focusing on the design of a hydrogen refuelling infrastructure for heavy goods vehicles. The spatio-temporal mapping is performed, within the demand data analysis component, for two time instances, a rollout phase and a mature phase, whereby the set of candidate refuelling sites and the refuelling demand for each site were identified. The optimisation model is developed within the second component of the framework and the optimal hydrogen refuelling infrastructure for heavy goods vehicles for Northern England, according to the given constraints and assumptions, is determined. A parameter variation and scenario analysis is further performed to investigate the maximum coverage distance parameter and the effect of having only localised (on-site) or only centralised (off-site) production supply of hydrogen. It is deduced that the system optimality is limited by the feasible number of tube trailer deliveries per day which suggests an opportunity for alternative delivery methods. Finally, an agent-based model is developed to simulate the operations of the refuelling infrastructure and analysis is performed on the sizing of the hydrogen storage capacity at hydrogen refuelling stations.

**Funding**

The research was partly funded by the UK Engineering and Physical Sciences Research Council through the Network-H2 grant for a network for hydrogen-refuelled transportation EP/S032134/1.
Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES


