Optimization Framework for HRS Location in North Netherlands

Research Methods for TOM
Dr. J.A. (José Alejandro) López Álvarez

Date: 10-06-2024

Authors (Group 2):
Guillermo Gil de Avalle - S5787084
Sareidakis Iosif- S57460878
Elisabeta Veliaj - S5792010
Maria Niki Antonopoulou- S5803047

Abstract: Hydrogen fuel cell electric vehicles (FCEVs) can reduce transportation emissions, but their adoption may be hindered by insufficient refuelling infrastructure. This study proposes a model for selecting current stations to be upgraded into hydrogen refuelling stations (HRS), in order to maximise coverage. By incorporating traffic data, the model moves beyond traditional approaches that rely on static factors to estimate refuelling station demand. Sensitivity analysis is incorporated to account for the effects of urban and highway areas, as well as different investment levels. The findings provide a framework for policymakers, suggesting a staggered approach to direct future investments toward greener transportation.

Table of Contents

Introduction

Hydrogen fuel cell electric vehicles (FCEVs) present a promising solution to reduce transportation emissions, but their adoption is limited by insufficient hydrogen refuelling infrastructure. Effective placement of hydrogen refuelling stations (HRS) is crucial to address this challenge. Traditional optimization models often focus on static factors such as geographical distance, minimising costs while ensuring coverage. However, they overlook dynamic factors like traffic intensity, which significantly impact HRS efficiency. Recent studies have begun to incorporate these dynamic elements. For example, Li et al. (2022) integrated traffic data into station planning to enhance network efficiency, while Chen et al. (2023) developed models considering traffic patterns and refuelling behaviour for more accurate HRS placement. Wang et al. (2023) highlighted the importance of data in improving the adaptability of hydrogen infrastructure. Building on this research, our study proposes an optimization model using traffic intensity data from sensors to maximise HRS coverage based on actual road usage. Sensitivity analyses on cost and distance parameters suggest that higher investment improves coverage, providing insights into HRS network efficiency and effectiveness. The findings aim to guide policymakers in strategic planning and infrastructure development, promoting hydrogen fuel technology adoption and supporting the transition to sustainable energy sources (Qadir et al., 2021; Apostolou & Xydis, 2019).

Research Gap

The main inspiration for this paper is the optimization study of Sikkema (2023), where a set of current gas stations was selected for an HRS upgrade, with and without on-site production capability, based on location and production of green supply facilities, as well as location of main cities. The objective of Sikkema's model was to guarantee coverage and production while minimising investment costs. However, no other parameters were used in order to estimate the actual traffic demands nodes, other than static representations of the cities locations. It also did not offer any estimation of the ROI for investment, other than how many cities could potentially be covered.

The paper from Xu et al. (2022) created a simulation that captures real traffic patterns and user equilibrium, offering a more accurate understanding of how travellers choose routes and improving station placement. However, no real demand data was used for said simulation, rendering it unable to actually locate actual demand nodes. Moreover, the simulation focused only on bus traffic in urban areas, which limits the conclusions of other areas and transportation type, e.g., freight transport in highways.

Finally, the paper from Wang et al. (2023) created a microsimulation that focused on determining suitability of battery electric vehicles (BEVs) and hydrogen fuel cell vehicles (HFCVs) in freight transportation, by optimising locations for refuelling and charging stations, and by establishing a framework to determine when the hydrogen freight vehicles were preferable to electric vehicles. Nevertheless, this study still missed actual traffic data in order to enable the calculation of demand.

As mentioned in the introduction, many papers have already suggested the inclusion of traffic intensity data for demand estimation purposes. However, due to the newness of hydrogen investigation for transport purposes and the difficulty of accessing and retrieving large quantities of data, few studies have actually utilised this approach. These reasons justify our research question, which states as follows:

Research Question:

How can the placement of hydrogen refuelling stations (HRS) be optimised to maximise network efficiency, considering dynamic demand factors like traffic intensity?

Methodology

To address the research question, the study proposes an optimization model that selects which existing gas stations can be upgraded to hydrogen refuelling stations (HRS) by integrating traffic intensity data that helps estimating refuelling demand nodes. The focus regions are Groningen, Fryslan, and Drenthe. Due to the time constraints of this study, only one day worth of traffic data is considered.

Data Description

Two primary data sources are utilised: traffic sensors data retrieved NDW (the Dutch National Traffic Data Portal), collected on May 7th, 2024 for that entire day, from their Dexter platform, and gas station locations from Overpass Turbo.

The traffic data, aggregated to obtain one daily observation per sensor, includes 1,548 observations after cleaning, with details on longitude, latitude, intensity, average intensity, sensor ID, and sensor downtime. This data captures actual road usage patterns and was cleaned to remove outliers and ensure consistency. The gas station data consists of 580 discrete locations, including longitude, latitude, and station ID.

Mathematical model

The mathematical model, detailed in Appendix D, aims to maximise the amount of cars that are served within a radius of any upgraded station. As such, our objective function considers a binary variable indicating whether any given station and sensor are within an acceptance distance range (z_{ik}), and the average traffic intensity -accounting for the downtime- for each sensor (I_k).

Objective Function

$$\text{Maximise } \sum_{i\in S}\sum_{k\in C} z_{ik} \cdot I_k$$

As mentioned, the aim is to quantify the cars passing in front of a sensor. However, since a car can pass several sensors on the same trip, 'cars served' should be considered a proxy for trips. To explain the correlation between this proxy and the studied variable, the model reports on "coverage," expected to be equal for both total trips and total cars served.

$$Cars\ served\ coverage\ (\%) \ = \frac{Cars\ served\ by\ the\ model\ in\ every\ instance}{Total\ number\ of\ cars\ detected\ by\ sensors}$$

Key constraints ensure each sensor is within a maximum acceptance radius from an HRS, no sensor is covered by more than one station, and total installation costs do not exceed the budget. These constraints ensure practical and cost-effective solutions.

Experimental design

Parameters

Several parameters are established in the model, which are crucial for optimising HRS placement. These include, amongst others, a straight-line distance calculation between stations and sensors (d_{ik}), the maximum distance a station can cover (D_{max}), and the average intensity at each sensor (I_k), which already accounts for the sensor downtime.

However, due to the impossibility to calculate the upgrade/installation cost per station (c_i), it was obtained from the literature. This cost includes necessary infrastructure for the establishment of six fuel dispensers (i.e., Fuel Dispensers, Canopy, Storage & Compression,and Site Preparation), excluding all other complementary costs, and averaging €3.85 million per station (Mayer et al., 2015), as detailed in Appendix E.

Sensitivity Analysis

The sensitivity analysis in this study evaluates the impact of two key parameters- total investment budget and maximum distance- on the placement and effectiveness of hydrogen refuelling stations.

Total Investment Budget

The total available budget (B) sets the financial limit for installations, ensuring practical and effective sensor coverage within budget constraints. However, since hydrogen as a road transportation fuel is a relatively new technology, it is difficult to find examples of individual budgets with this purpose. For that reason, and due to the scientific relevance of observing the outcomes of different investment levels, we propose a sensitivity analysis for total investment budget, with 10 iterations ranging from €20 million to €100 million, as depicted in Appendix C.

Maximum Distance

Keeping HRS within a convenient range boosts FCEVs adoption (Kuby et al., 2009). Economic viability is also important to consider, as stations too far apart may be underused, while those too close together incur extra costs and competition (Kim et al., 2021). This paper acknowledges that lower distances between sensors (i.e., cars served) and stations will likely be more suitable for urban areas, where drivers are not willing to drive long distances for refuelling, whilst long distances will likely suit more highway transportation, where the potential detour enlarges. Since it is increasingly difficult to control for drivers' detour desires in uncontrolled environments, we propose a sensitivity analysis for total maximum distance, with 5 iterations ranging from 1Km to 5Km, as depicted in Appendix C.

Results and Discussion

Experiment settings

The model runs in Python, using the Gurobipy to run the optimization model, which is a renowned library with different available optimization algorithms (Simplex, Monte Carlo, etc.). Since the data was pre-cleaned, the model is directly loaded. However, in order to account for the redundancy of such a large number of sensors, a threshold is established, filtering out the 25% with lowest intensity. Thereafter, distance between sensor and station is pre-calculated using a vectorized calculation for efficiency, and the model is iterated using a nested for loop, in order to account for the different iterations within the sensitivity variables.

Results presentation

The results presented 50 iterations for the levels within the ranges set for the two sensitive variables, i.e., Maximum distance, and Total investment. Detailed results, available in Appendix C, include the total cars served per hour/day, daily percentual coverage, number of stations upgraded, and average cost per customer.

However, it’s crucial to first understand the sensors with the highest traffic density. Exploring the correlation between these sensors and the upgraded stations can reveal patterns that inform decision-making. Figure 1 shows a heatmap of all sensors (left) and a map of the top 10% highest demand sensors (right). The areas of greatest intensity are mainly Groningen (main city), as well as Heerenveen and Meppel, pivotal locations on the route to the Randstad area.

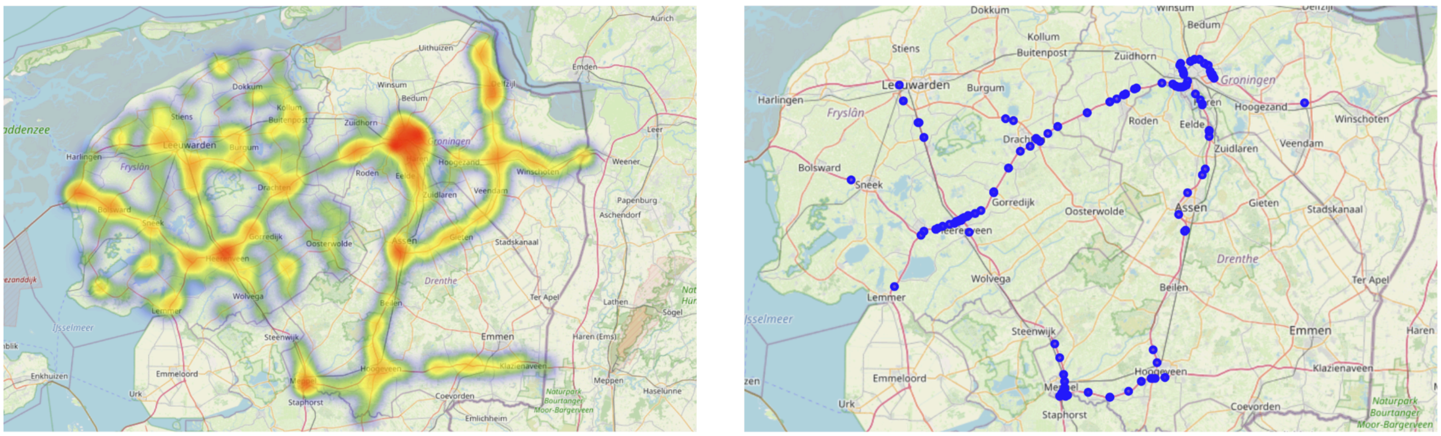

Figure 1: Analysis of the most commonly served areas (heatmap on the left; locations with the top 10% highest demand on the right).

Subsequently, we proceed with the analysis of the model results, initially focusing on the 'Max distance' sensitivity analysis. To simplify the results' interpretation, Figure 2 displays three maps produced by the model, where the 'Max distance' parameter iterates from 1km to 5km to 10km. Using this progressive approach, with a fixed cost of 55 million, we can examine the effects of distance constraints on station placement and coverage.

As depicted on the maps, the number of upgraded stations remains constant at 14 across all scenarios due to the fixed investment, suggesting a linear correlation between investment and station numbers. However, coverage varies drastically, from 40% in the first scenario to 94% in the last. Initially, stations are concentrated in the previously identified high-traffic (Groningen, Heerenveen, and Meppel). As the distance boundary increases, without a constraint that limits a station's capacity, one station can cover all high-intensity areas, allowing redistribution of the rest of stations to medium and low-demand sensors. Beyond a certain distance, coverage suffices to reach most sensors, indicating a parabolic relationship (i.e., beyond a certain point, additional stations no longer increase coverage).

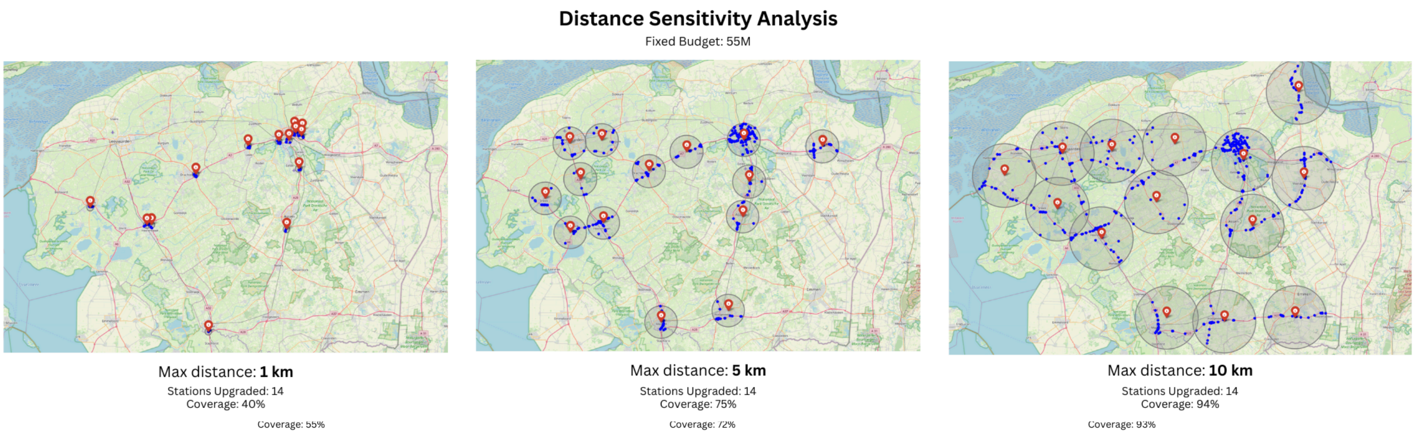

Figure 2: Maximum Distance Sensitivity Analysis for a fixed "Total Cost" of 55M

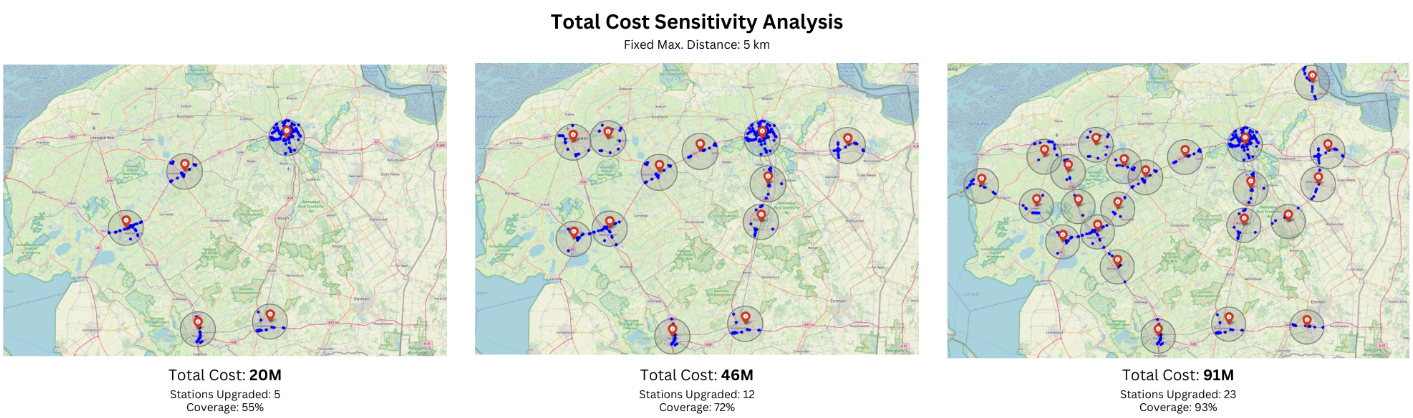

Figure 3: Total Cost/Investment Sensitivity Analysis for a fixed "Max distance" of 5km

A similar analysis is conducted for the total investment sensitivity, where the maximum distance is fixed at 5 km and investment levels are varied from 20M to 46M to 91M, covering most of the studied range. Figure 3 displays three maps. We observe that, in this scenario, both station upgrades and coverage increase linearly, confirming our previous assumption for the correlation between budget and coverage. In terms of location, upgrades initially focus on areas of high intensity. As the budget increases, they gradually extend to areas of medium intensity, further reinforcing the linear relationship.

Figure 4 illustrates coverage for different "Max Distances" in different budget levels. This figure further confirms the linear relationship between budget and coverage for smaller distance boundaries, which turns to parabolic functions as the distances are increased. Since smaller distances are more suitable for urban areas, and larger distances for highways, we can infer that budget-coverage linearity will be more present in cities, and parabolic correlations will be found mostly in highways.

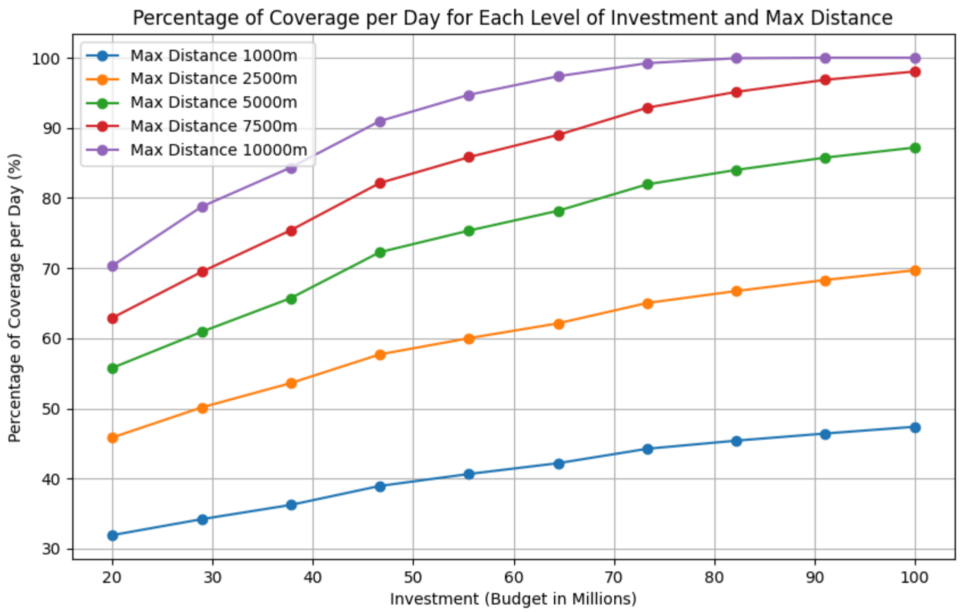

Figure 4: Daily coverage (%) for various "Max Distances" in different levels of Total Investment

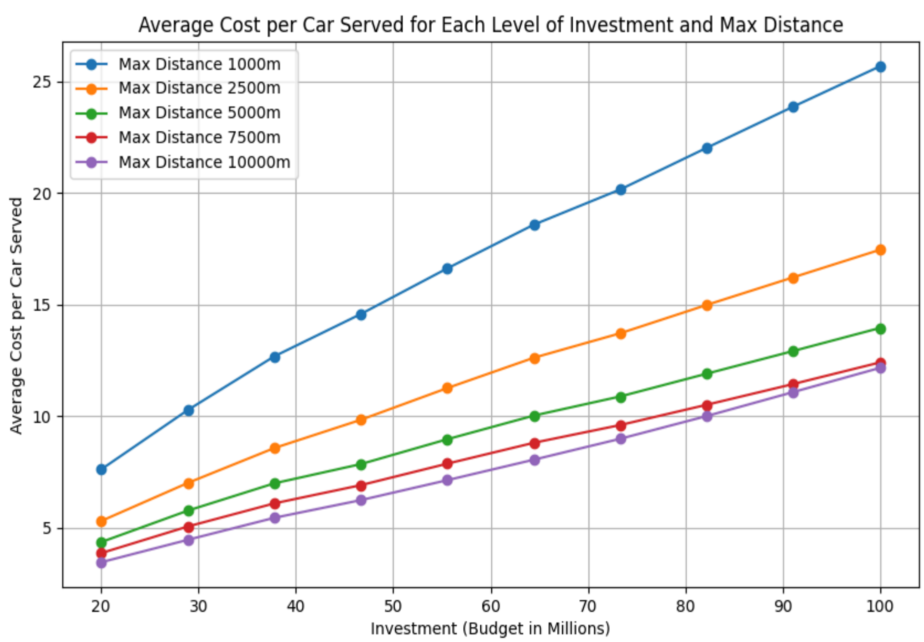

Figure 5: Average cost per "opportunities" served for different configurations.

Next, we correlated the total cars covered with the total investment, yielding an average cost per car served. Figure 5 illustrates the results for this analysis. At smaller distances, there is a sharper and more pronounced increase in average costs, while at greater distances, we observe a smoother and exponential decrease in costs across all investment levels. This suggests three key insights: 1) Financially, it is more advantageous to develop larger distances. 2) When targeting smaller distances, such as in urban areas, it is more effective to make smaller investments that cover high-demand sensors. 3) Beyond a certain point, further increasing the distance becomes ineffective, indicating that there is a limit to the cost benefits of expanding coverage.

Finally, a comparison between the HRS average cost per customer and other popular fuels is offered in Appendix E. Figure 6 shows the outcomes for said comparison, where we observe that the average price per customer for hydrogen almost quadrupled the next most expensive fuel, positioning itself as the most expensive fuel per initial costs. Despite this, hydrogen storage cost is much lower than other green fuels (such electric vehicles), which can offset the initial cost, specially for transportation modes where high storage capacity is important -like freight- which confirms our suggestion that HRS are more financially attractive for longer distances (i.e., highways).

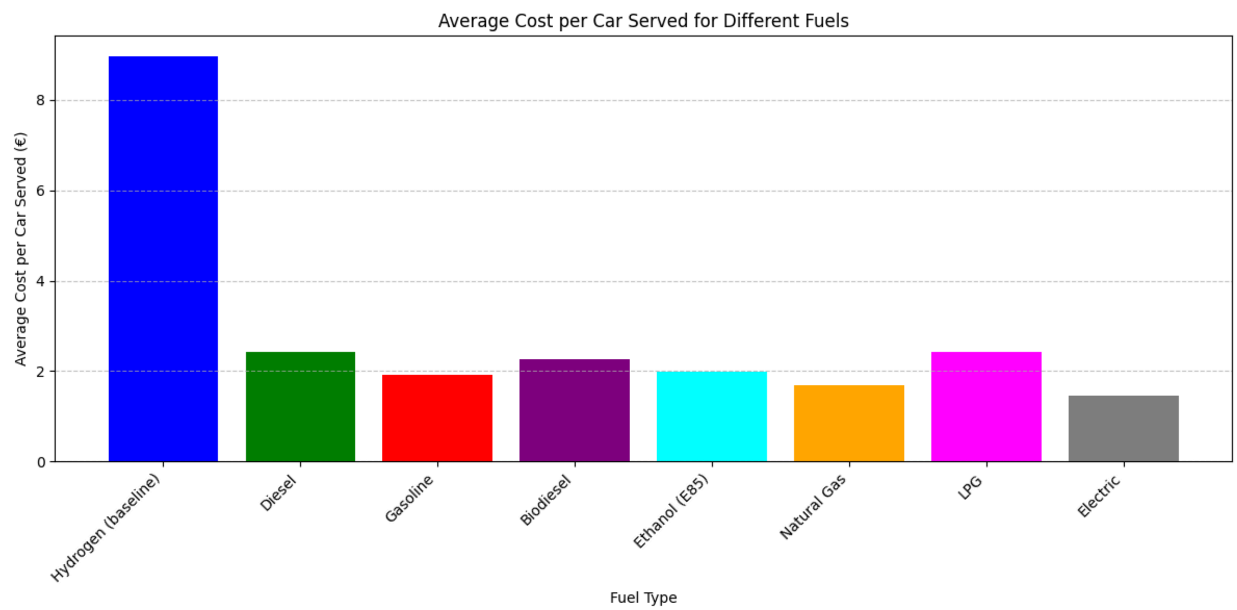

Figure 6: Average cost per car served for different fuel types

Discussion of Insights

The aforementioned findings suggest a staggered approach for investment prioritisation, where investments are staged according to their returns in terms of coverage and cost, a common practice in large investments in the sector.

The initial stage should focus on developing HRS along highways with medium-high demand, targeting a station coverage distance of 5 - 7.5 km. This stage would ensure a very high level of coverage on the highway areas due to the parabolic distribution for larger distances -which should specially benefit freight transport-, while maintaining a low and linear average cost per potential customer.

The second stage should concentrate on urban areas with high demand. Even at shorter station coverage distances, modest investments targeted at the studied high-demand points could yield substantial coverage impacts. Additionally, the moderate nature of the investment helps keep the average cost per customer low, as the returns on coverage diminish once these points are addressed.

Conclusion

The outcomes from our model, which integrates logical constraints with traffic density data, suggests a framework for formulating financially efficient investment strategies for HRS infrastructure in similar regions. This framework not only pinpoints locations for new stations but also provides proxy measurements for estimating the investment's ROI, such as coverage and average cost per customer. Moreover, it recommends a staggered investment approach, where different indications are given around how to stage the investments, which areas to prioritise, and how to understand the correlations between areas, budgets, and traffic intensity.

Additionally, this study's findings both align with and diverge from those of Sikkema (2023). While Sikkema's approach focuses on cost reduction to minimise investment, our approach prioritises maximising vehicle coverage and scalability to inform investment decisions. Both papers propose a set of stations for upgrade, but our study also offers additional investment estimators such as ROI and coverage, which are relevant for stakeholders.

In conclusion, this framework can assist policymakers and relevant organisations (e.g., gas companies) in strategizing, shaping policies, and making informed decisions related to the hydrogen 'chicken or egg' dilemma, which questions whether the low adoption rate of FCEVs or the current HRS capacity is the primary issue. A comprehensive set of limitations and future research is offered in Appendix F.

References

Apostolou, D., & Xydis, G. (2019). A literature review on hydrogen refuelling stations and infrastructure. Current status and future prospects. Renewable and Sustainable Energy Reviews, 113, 109292. https://doi.org/10.1016/j.rser.2019.109292.

Chen, Y.H., Zhang, W., & Li, S. (2023). Energy management optimization of fuel cell hybrid ships based on particle swarm optimization.

Kuby, M., Lines, L., Schultz, R., Xie, Z., Kim, J. G., & Lim, S. (2009). Optimization of hydrogen stations in Florida using the Flow-Refueling Location Model.

Leaflet. (n.d.). Overpass Turbo. Retrieved May 2024, from https://overpass-turbo.eu/

Li, S., Xu, M., Huang, W., & Qin, Z. (2022). Finite Element Analysis and Optimization of Hydrogen Fuel Cell City Bus Body Frame Structure.

Mayer, T., Semmel, M., Morales, M.A.G., & Schmidt, K.M. (2015). Techno-economic evaluation of hydrogen refuelling stations with liquid or gaseous stored hydrogen.

NDW. (n.d.). Dexter: Data for transportation professionals. Retrieved May 2024, from https://www.ndw.nu/dexter-platform/

Qadir, S. A., Al-Motairi, H., Tahir, F., & Al-Fagih, L. (2021). Incentives and strategies for financing the renewable energy transition: A review. Energy Reports, 7, 3590-3606. https://doi.org/10.1016/j.egyr.2021.06.041

Sikkema, P. (2023). Hydrogen Refuelling Infrastructure Expansion in the Netherlands: A Set Covering Model Considering On-Site Hydrogen Production and Appropriability for Specific Locations. University of Groningen, Faculty of Economics and Business.

StartupsMaker. (2024). How much does a gas station cost?. Retrieved May 2024, from https://startupsmaker.com/how-much-does-a-gas-station-cost/#Fuel_Dispensers

Wang, S., Peng, Z., Wang, P., Chen, A., & Zhuge, C. (2023). A data-driven multi-objective optimization framework for determining the suitability of hydrogen fuel cell vehicles in freight transport. International Journal of Hydrogen Energy, 48(12), 4567-4580. https://doi.org/10.1016/j.ijhydene.2023.01.056

Xu, T., Li, L., & Fan, S. (2022). Hydrogen station allocation based on equilibrium traffic flow. International Journal of Hydrogen Energy, 47(4), 3456-3468.

Alternative Fuels Data Center. (n.d.). Vehicle technologies office. U.S. Department of Energy. Retrieved June 2024, from https://afdc.energy.gov

Kim, H., Kim, B.-I., & Thiel, D. (2021). Exact algorithms for incremental deployment of hydrogen refuelling stations. International Journal of Hydrogen Energy, 46(56), 28760-28774. https://doi.org/10.1016/j.ijhydene.2021.06.104

Appendix A: Response to Reviews

This appendix details our responses to feedback received by Group 6 during the development of our study "Optimization Framework for HRS Location in North Netherlands."

- **Title:**

  Comment: The title provides a positive and optimistic view of the hydrogen problem but it could be more specific, mentioning the exploration of alternative energies and the methodology.

  Response: We acknowledged that the specificity could be improved. Therefore, we revised the title to: “*Optimization Framework for HRS Location in North Netherlands”.*

- **Motivation:**

  Comment: The motivation is clear but too broad for the audience. The focus should be more specific to the research question.

  Response: We took into consideration both the comment from the other team and the feedback from our coordinator and we revised our research proposal to adopt a more scientific tone including more detailed background literature and motivation behind our study. Furthermore, we broadened our research question to have global relevance, focusing on how traffic intensity affects HRS location optimization, which can be applicable to various regions beyond the Netherlands.

- **Research Question:**

  Comment: The research question is broad with many proposals, A more focused approach is needed.

  Response: We refined our research question to focus specifically on how traffic intensity data can optimise HRS placement. Therefore, we suggest the revised research question: “*How can the placement of hydrogen refuelling stations (HRS) be optimised under budget constraints to maximise network efficiency, considering traffic intensity, and installation costs?”.* This narrower scope helps us effectively to tackle the study within the given timeframe.

- **Methodology:**

  Comment: The methodology is adequate but needs a clearer explanation, particularly regarding the optimization model and the methodology in mathematical problems.

Response: We revised the methodology section to provide a detailed explanation of our optimization model. In terms of a mathematical model, we incorporated the new decision variable Zik , indicating whether a facility is serving the demand of a specific sensor, to refine our model's precision. Additionally, we added constraints to ensure practical applicability, such as station capacity and distance limitations, enhancing the model's relevance.

- **Experimental Design:**

  Comment: The section needs to be separated and more detailed, especially regarding the three scenarios for the gas station location.

  Response: We created a separate experimental design section that details the steps, datasets and specific scenarios for gas station locations. We included a detailed report on the sensors' locations and their coverage on maps, providing a visual representation of our findings. We also documented the stations that consistently appeared across different budget and distance scenarios, highlighting key sites for HRS deployment. As a final step, we identified and reported the sensors with the highest coverage and demand, ensuring focused investment in high-impact areas.

- **Scientific Contribution:**

  Comment: The focus is too broad and there is a lack of data sources and evidence.

  Response: We narrowed our focus to the impact of traffic intercity on HRS placement. Additionally, we have included more data sources and evidence in our revised sections, ensuring that our scientific contribution is both solid and well-supported by data.

Appendix B: Extensive Results

Run	Max Distance (metres)	Budget (€)	Total cars per hour	Coverage (%)	Total cars per day	Average cost per car served	Number of stations upgraded
1	1000	20M	109,447	31.94	2,626,728	€7.61	5
2	1000	28.8M	117,209	34.21	2,813,016	€10.27	7
3	1000	37.7M	124,182	36.24	2,980,368	€12.68	9
4	1000	46.6M	133,457	38.95	3,202,968	€14.57	12
5	1000	55.5M	139,309	40.66	3,343,416	€16.62	14

6	1000	64.4M	146,888	42.88	3,525,312	€18.57	16
7	1000	73.3M	154,744	45.13	3,713,856	€20.15	19
8	1000	82.2M	158,513	46.23	3,804,312	€22.02	21
9	1000	91.1M	164,689	47.8	3,952,536	€23.86	23
10	1000	100M	169,123	49.08	4,059,648	€25.66	25
11	2500	20M	157127	45.859	3771048	€5.30	5
12	2500	28.8M	171797	50.141	4123128	€7.01	7
13	2500	37.7M	183689	53.611	4408536	€8.57	9
14	2500	46.6M	197717	57.706	4745208	€9.83	12
15	2500	55.5M	205641	60.018	4935384	€11.26	14
16	2500	64.4M	212865	62.127	5108760	€12.61	16
17	2500	73.3M	222841	65.038	5348184	€13.71	19
18	2500	82.2M	228669	66.739	5488056	€14.98	21
19	2500	91.1M	234077	68.318	5617848	€16.22	23
20	2500	100M	238794	69.694	5731056	€17.45	25
21	5000	20M	191145	55.787	4587480	€4.36	5
22	5000	28.8M	208719	60.917	5009256	€5.77	7
23	5000	37.7M	225175	65.719	5404200	€6.99	9
24	5000	46.6M	247644	72.277	5943456	€7.85	12
25	5000	55.5M	258186	75.354	6196464	€8.97	14
26	5000	64.4M	267862	78.178	6428688	€10.02	16
27	5000	73.3M	280772	81.946	6738528	€10.88	19
28	5000	82.2M	287829	84.006	6907896	€11.90	21
29	5000	91.1M	293851	85.763	7052424	€12.92	23
30	5000	100M	298737	87.189	7169688	€13.95	25

31	7500	20M	215585	62.92	5174040	€3.87	5
32	7500	28.8M	238008	69.465	5712192	€5.06	7
33	7500	37.7M	258344	75.4	6200256	€6.09	9
34	7500	46.6M	281437	82.14	6754488	€6.91	12
35	7500	55.5M	294063	85.825	7057512	€7.87	14
36	7500	64.4M	304906	88.99	7317744	€8.81	16
37	7500	73.3M	318205	92.871	7636920	€9.60	19
38	7500	82.2M	325946	95.13	7822704	€10.51	21
39	7500	91.1M	331850	96.853	7964400	€11.44	23
40	7500	100M	335877	98.029	8061048	€12.41	25
41	10000	20M	241060	70.356	5785440	€3.46	5
42	10000	28.8M	269820	78.749	6475680	€4.46	7
43	10000	37.7M	288985	84.343	6935640	€5.45	9
44	10000	46.6M	311672	90.964	7480128	€6.24	12
45	10000	55.5M	324506	94.71	7788144	€7.13	14
46	10000	64.4M	333573	97.356	8005752	€8.05	16
47	10000	73.3M	339960	99.22	8159040	€8.99	19
48	10000	82.2M	342385	99.928	8217240	€10.01	21
49	10000	91.1M	342631	100	8223144	€11.08	23
50	10000	100M	342631	100	8223144	€12.16	23

Appendix C: Mathematical Model

Sets and Indices

- S: Set of potential stations, indexed by *i*.
- C: Set of sensors, indexed by *k*.

Parameters

- d_{ik}: Distance (as a straight line) between station i and sensor *k*.
- *D*: Maximum allowable distance for a station to cover a sensor (sensitive variable).
- I_k: Average demand per hour at sensor *k*.
- c_i: Installation cost for station *i*.
- *B*: Total available budget for installing stations (sensitive variable).

Decision Variables

- $x_i \in \{0, 1\}$: Binary variable indicating whether to place a station at location *i*.
- $z_{ik} \in \{0, 1\}$: Binary variable indicating whether station i covers sensor *k*.

Objective Function

Maximise the total coverage of sensors weighted by their hourly demand:

$$\text{Maximise } \sum_{i \in S} \sum_{k \in C} z_{ik} \cdot I_k$$

Constraints

1. Distance Constraint: A sensor can only be covered by a station within

the maximum distance.

$$z_{ik} \leq y_{ik} \cdot x_i \quad \forall\, i \in S, \ \forall\, k \in C$$

where $y_{ik} = 1\ if\ d_{ik} \leq D$, otherwise $y_{ik} = 0$.

2. Sensor Coverage Constraint: Each sensor can be covered by at most one station.

$$\sum_{i \in S} z_{ik} \leq 1 \quad \forall k \in C$$

3. Budget Constraint: The total installation cost cannot exceed the available budget.

$$\sum_{i \in S} c_i \cdot x_i \leq B$$

4. Binary Variables:

$$x_i \in \{0, 1\} \quad \forall i \in S$$

$$z_{ik} \in \{0, 1\} \quad \forall i \in S, \ \forall k \in C$$

Appendix D: Comparison with other fuel costs

To establish a comparison between hydrogen costs in our model and other fuels, we need to first define the costs to consider and estimate them for the most common fuels. Since our aim is not to estimate the cost of a complete gas station (which includes typical additional services), we will focus solely on the cost of essential equipment and construction for six fuel dispensers (or charging stations). These typically account for 30-40% of the total cost (StartupsMaker, 2024). For each instance, we will assume that all six dispensers installed are of the same fuel type.

Fuel Type	Fuel Dispensers	Canopy	Storage & Compression	Site Preparation	Total Cost Estimate	Average Cost Estimate
Hydrogen (baseline)	€1,185,000 - €2,400,000	€80,000 - €100,000	€1,085,000 - €2,000,000	€350,000 - €500,000	€2,700,000 - €5,000,000	**€3,850,000**
Diesel	€90,000 - €240,000	€15,000 - €100,000	€300,000 - €1,200,000	€50,000 - €150,000	€455,000 - €1,690,000	**€1,072,500**
Gasoline	€60,000 - €180,000	€15,000 - €100,000	€240,000 - €900,000	€50,000 - €150,000	€365,000 - €1,330,000	**€847,500**
Biodiesel	€120,000 - €300,000	€20,000 - €100,000	€300,000 - €900,000	€50,000 - €200,000	€490,000 - €1,500,000	**€995,000**
Ethanol (E85)	€60,000 - €180,000	€15,000 - €100,000	€300,000 - €900,000	€50,000 - €150,000	€425,000 - €1,330,000	**€877,500**
Natural Gas	€120,000 - €300,000	€15,000 - €100,000	€250,000 - €500,000	€50,000 - €150,000	€435,000 - €1,050,000	**€742,500**
LPG	€90,000 - €240,000	€15,000 - €100,000	€300,000 - €1,200,000	€50,000 - €150,000	€455,000 - €1,690,000	**€1,072,500**
Electric	€12,000 - €600,000	€20,000 - €100,000	€50,000 - €300,000	€50,000 - €150,000	€132,000 - €1,150,000	**€641,000**

The table above breaks down an estimation of these costs, including the fuel dispensers themselves, the canopy to protect fuel from weather conditions, the storage and compression for the fuel (including the cost of an Underground Storage Tank, when relevant), and the necessary preparation of the terrain for the safe installation of energy infrastructure. The data was retrieved in June 2024 from the Alternative Fuels Data Center of the U.S. Department of Energy (n.d.).

In order to compare these costs with the coverage levels estimated by our model, we pick one of the iterations of the total budget and maximum distance to build a case for how the rest of the fuels would perform. Specifically, we select iteration 25, with a total budget of €55M and a maximum distance of 5 km between the station and sensor. This iteration covered approximately 75% of the sensors, accounting for a total of 6,196,464 cars with 14 stations updated. With these results at hand, the following average cost per customer was estimated:

Fuel Type	Average Cost Estimate	Multiplied by 14 stations	Average cost per car served
Hydrogen (baseline)	€3,850,000	€53,900,000	€8.97

Diesel	€1,072,500	€15,015,000	€2.42
Gasoline	€847,500	€11,865,000	€1.91
Biodiesel	€995,000	€13,930,000	€2.25
Ethanol (E85)	€877,500	€12,285,000	€1.98
Natural Gas	€742,500	€10,395,000	€1.68
LPG	€1,072,500	€15,015,000	€2.42
Electric	€641,000	€8,974,000	€1.45

Appendix E: Limitations and Future Research

Limitations

This study's optimization model has several limitations that may affect the bias and variance of its results. Firstly, the data is confined to the available NDW sensors, which are presumed to be concentrated in high-traffic areas. Moreover, due to time constraints, the analysis was limited to a single day of data (May 7th, 2024). Additionally, distances between stations and sensors were calculated as straight lines, not accounting for actual road patterns.

Furthermore, the model presupposes uncapped station capacity due to the current low adoption rate of hydrogen vehicles, allowing for an unlimited number of hydrogen cars to refuel at each station. This assumption might not hold as hydrogen vehicle adoption increases. Finally, the static nature of the traffic data, which can vary daily and seasonally, also presents a limitation. The model does not account for potential changes in traffic intensity over time, which could affect its long-term applicability.

To overcome these limitations, several measures could be taken. Expanding the data collection period to include multiple days or weeks would provide a more comprehensive dataset, capturing traffic patterns. Utilising advanced data aggregation and variance reduction techniques, such as Random Forest, could improve the robustness of the traffic data. Incorporating actual road network data for distance calculations would enhance the accuracy of the coverage results. Additionally, adjusting the model to include potential increases in hydrogen vehicle adoption and dynamic traffic patterns would improve its applicability over time.

Future Research

Future research could explore ideas that overcome the limitations of the study, such as adding more days of data to improve the robustness of the analysis. Incorporating software that includes road patterns, rather than straight-line distances, will enhance accuracy. Including station capacity by estimating the percentage of cars needing to refuel is essential to future-proof the model for different levels of adoption.

Additionally, other lines of research could include incorporating constraints such as a minimum distance between upgraded stations, which is common in this type of study. Another addition to the model could be geodata from hydrogen vehicle sales, which could further complement the demand estimation. Finally, expanding the model to other regions with different traffic patterns and needs could provide a holistic view of nationwide investment, ensuring broader applicability and relevance.