

# **A Fuel-Travel-Back Approach to Hydrogen Station Siting**

**Zhenhong Lin<sup>\*</sup>, Joan Ogden, Yueyue Fan, Chien-Wei Chen**

***University of California Davis***

Institute of Transportation Studies

One Shields Avenue

Davis, CA 95616

The United States

## **ABSTRACT**

The problem of hydrogen station location is often studied through understanding refueling behavior or reviewing the experience of gasoline stations. Driven by the notion "where you drive more is where you more likely need refueling", this paper develops a new approach where station siting is treated as a fuel-travel-back problem and the only required data is VMT distribution. Such a fuel-travel-back problem is a typical transportation problem and is solved as mix-integer-programming model. When the total fuel-travel-back time is minimized, so is the average refueling travel time of a random motorist, for which theoretical deduction is provided. The model is applied to derive an optimal station roll-out scheme for Southern California. The results show that, if station size constraints are relaxed, only 18% of existing gas station number is needed to achieve the current fuel accessibility of gasoline in the region. Fewer stations lead to larger station size, suggesting a need to re-examine the current speculation on designs of hydrogen station and distribution system and to conduct more regional studies for discovery of optimistic and pessimistic regions for hydrogen. The results also indicate that early stations should be located strategically and even at low-demand locations, which is contradictory to existing proposition.

Keyword: hydrogen; station location; alternative fuel; optimization

---

\* Corresponding author: [zlin@ucdavis.edu](mailto:zlin@ucdavis.edu) (Zhenhong Lin)

## **NOMENCLATURE**

ARTT = average refueling travel time

FCV = fuel cell vehicle

H<sub>2</sub> = hydrogen

VMT = vehicle miles traveled

## **1. INTRODUCTION**

Hydrogen as vehicle fuel offers the promise of reducing air pollution, greenhouse gas emissions, and oil dependence [1]-[3]. To introduce hydrogen vehicles, a refueling network must be built to ensure some certain level of fuel accessibility [4]-[6], but fuel accessibility could be extremely costly in an early market due to a high ratio of capital investment to demand. This raises the question of how to effectively locate a limited number of stations.

By examining alternative fuel experiences in the United States [4][5][7], New Zealand [8], Canada [9], these studies attempt to estimate a sufficient number of stations, in terms of percentages of existing gasoline stations, for a successful alternative fuel vehicle fleet, but do not explicitly consider where to locate the stations. In the field of operations research, the station siting problem has often been treated as facility location on a network of roads [10]-[13] or as a subset of the existing gasoline station network [14]. These facility location models explicitly or implicitly assume the home or workplace as the origin of refueling travel [10]-[14]. Berman [15] questions this origin assumption by pointing out that refueling is often a secondary purpose of travel, and proposes a flow-capture model. However, the major drawbacks of the flow-capture approach are ignorance of the difference of capturing long and short trips and ignorance of the inconvenience suffered by un-captured flows. It also requires origin-destination data, which are often difficult to obtain.

Driven by a mobile-origin notion that "where you drive more is where you more likely need refueling", this paper develops a fuel-travel-back approach that only requires data of *VMT* spatial distribution. Instead of the home or workplace, any point along the road network is a possible origin of the refueling trip, with the probability quantified by the distribution of *VMT* or fuel consumption. Then, station siting is treated as a network transportation problem where the burned fuel along the road hypothetically travels back to the nearest stations and the objective is to minimize the total fuel-travel-back travel time. Some practical issues of the model are discussed, followed by results and discussions on the Southern California case study.

## 2. METHOD AND DATA

### 2.1. Fuel accessibility

Fuel accessibility is defined as the easiness for a random motorist to access a station from where the motorist has a refueling need. This section addresses three underlying questions:

- Who is this random motorist?
- Where are the origins of refueling trip?
- How the easiness is measured?

Consider a directed graph,  $G = (N, A)$  with node set  $N$  and arc set  $A$ . Let  $i, j \in N$  be any node  $i$  and  $j$ , and  $a_{i,j} \in A$  be a directed arc from node  $i$  to  $j$ . Let  $s = s_{a,m\Delta} \in S$  denote a small segment on arc  $a_{i,j}$ , with a fixed small length of  $\Delta$ , and with a distance of  $m \cdot \Delta$  from node  $i$  ( $m = 1, 2, \dots$ ). Set  $S$  contains all the small segments on all arcs. Let  $V$  denote the set of all motorists traveling along  $G$  and  $v \in V$  be any particular one of them.

At this point, one node can at most have one station, although multiple stations at one node will be discussed later. Let  $N^f \in N$  be the set of refueling nodes. Now consider a random motorist  $v^* \in V$  driving on  $G$  to assess the fuel accessibility of  $N^f$ . One proper measurement of fuel accessibility of  $N^f$  is the expected value of travel time for  $v^*$  to travel from where  $v^*$  need a

refueling to the nearest station, defined as average refueling travel time ( $ARTT$ ). There are two sources of randomness here:  $p^v$ , the probability of  $v^*$  being any particular motorist  $v$ ; and  $p^{vs}$ , the probability of this particular motorist  $v$  having a refueling need at a particular location  $s$ . Once these two probabilities as well as the travel time from  $s$  to the nearest station, denoted as  $t^s = t^s(N^f)$ , can be quantified,  $ARTT$  can be calculated via equation (1).

$$ARTT = \sum_{v \in V} \sum_{s \in S} t^s \cdot p^v \cdot p^{vs} \quad (1)$$

Both  $p^v$  and  $p^{vs}$  are independent of  $N^f$  and need further formulation, for which some terms need to be defined. For simplicity, a time frame of one year is assumed. Let  $f^{vs}$  be the number of times per year  $v$  passing  $s$ ,  $f^v$  be the total number of visits by  $v$  to everywhere in  $S$ ,  $f^s$  be the total number of all motorists  $V$  to a specific location  $s$ , and  $f$  be the total number of visits by all motorists  $V$  to everywhere in  $S$ .

The attributes of location  $s$  that contribute to a large  $f^{vs}$  could be closeness to  $v$ 's home, workplace, or favorite shopping center, or just belonging to  $v$ 's most enjoyable route. Whatever the reasons,  $f^{vs}$  aggregately reflects  $v$ 's travel behavior caused by the network, perception, budget, etc. Intuitively, where one drives more is where one more likely needs refueling, so a larger  $f^{vs}$  implies a larger  $p^{vs}$ , which, as an assumption, is represented by equation (2). When  $v$  at  $s$  has a refueling need,  $s$  becomes the origin of the refueling trip.

$$p^{vs} = f^{vs} / f^v \quad (2)$$

As another assumption, the probability of  $v^*$  being a particular motorist  $v$  is weighed by  $v$ 's relative travel frequency, as in equation (3). The implication of this assumption is that more frequent drivers have more votes on deciding where stations should be located.

$$p^v = f^v / f \quad (3)$$

Combining equations (1) through (3), we can obtain another form of  $ARTT$  as in equation (4), which is important in that the motorist index  $v$  disappears and therefore disaggregate travel data, difficult to obtain, become unnecessary.

$$\begin{aligned} ARTT &= \sum_{v \in \mathcal{V}} \sum_{s \in \mathcal{S}} t^s \cdot f^{vs} / f \\ &= \sum_{s \in \mathcal{S}} t^s \cdot \sum_{v \in \mathcal{V}} f^{vs} / f = \sum_{s \in \mathcal{S}} t^s \cdot f^s / f \end{aligned} \quad (4)$$

Equation (4) is equivalent to either equation (5) or (6), where  $T^s$  and  $T$  represents VMT at location  $s$  and the whole network  $S$ , and  $FUEL^s$  and  $FUEL$  are the corresponding fuel consumption.  $\Delta$ , as previously defined, is the length of any location  $s$ .  $C_{fe}$  represents fuel economy. Equation (5) indicates that the only needed data is the spatial distribution of VMT, which is usually not difficult to obtain. Equation (6) and (7) describe an interesting theorem: minimizing average refueling travel time is equivalent to minimizing the total travel time for the fuel to travel from where it is burned back to the nearest station. So station siting is equivalently transformed into the fuel-travel-back context. It is merely a hypothetical analogy to aid communication, as fuel can not be transported back once it is burned.

$$ARTT = \sum_{s \in \mathcal{S}} t^s \cdot \frac{f^s \cdot \Delta}{f \cdot \Delta} = \sum_{s \in \mathcal{S}} t^s \cdot \frac{T^s}{T} \quad (5)$$

$$ARTT = \sum_{s \in \mathcal{S}} t^s \cdot \frac{T^s / C_{fe}}{T / C_{fe}} = \sum_{s \in \mathcal{S}} t^s \cdot \frac{FUEL^s}{FUEL} \quad (6)$$

$$ARTT \cdot FUEL = \sum_{s \in \mathcal{S}} t^s \cdot FUEL^s \quad (7)$$

## 2.2. Stations Siting as a Transportation Problem

The fuel-travel-back perspective of equation (7) establishes a perfect context for explaining our optimization model. Apparently, the objective is to minimize the total time for fuel-travel-back. Hypothetically, if stations are built everywhere, then any  $t^s$  is zero and there is no travel time. The problem is to minimize the total time for the constraint of station number  $|N^f|$ .

For all directional road segments pointing to node  $j$ , the fuel quantities  $FUEL^s$  along these segments first arrive at node  $j$ , gather into an aggregate demand  $FUEL^j$ , and "look for" the nearest station. If we introduce  $l^j$  as the average time for these  $FUEL^s$  to travel from  $s$  to node  $j$ , the problem can then be formulated as a typical transportation problem [16], as in equation (8). There are two decision variables:  $flow^{ji}$  representing the amount of fuel traveling from  $j$  to  $i$ ;  $build^i$  representing whether or not to build a station at node  $i$ . Constraint (a) ensures satisfaction of all demands, as like forcing all fuel to travel back to stations. Constraint (b) ensures that fuel can only travel back to a refueling node and  $Mnum$  is an arbitrarily big number merely for programming purpose. Constraint (c) limits the number of refueling nodes to be  $|N^f|$ .

Minimize:

$$FUEL \cdot ARTT = \sum_{j,i \in N} (t^{ji} + l^j) \cdot flow^{ji}$$

Subject to:

$$\sum_{i \in N} flow^{ji} = FUEL^j \quad (\forall j \in N) \quad (a)$$

$$\sum_{j \in N} flow^{ji} \leq Mnum \cdot build^i \quad (\forall i \in N) \quad (b) \quad (8)$$

$$\sum_{i \in N} build^i = |N^f| \quad (c)$$

$$build^i = \begin{cases} 1 & i \text{ is refueling node} \\ 0 & \text{otherwise} \end{cases}$$

### 2.3. Some Practical Issues

- Location continuity

In practice, we are more interested in how a refueling network grows rather than just a static station siting scheme. The term  $|N^f|$  in equation (8) represents the total number of refueling nodes. If we set  $|N^f| = 1, 2, \dots$  and apply the model independently for each  $|N^f|$ , we can obtain a station roll-out scheme consisting of a series of static station location schemes. However, by adopting such a roll-out scheme to describe the growth of refueling network, we are ignoring the

spatial relationships among these static schemes or assuming it is free of cost to move a station from one location to another, either of which is inappropriate. This is referred to as the location continuity issue.

The location continuity issue is handled by posing the constraint of location subset. That is, the optimal locations of  $|N^f|$  stations must be a subset of those of  $|N^f| + 1$  stations. This ensures the refueling network grows logically.

- Starting number of refueling nodes

Related to the location continuity issue is the issue of starting number of refueling nodes. It is about how many refueling nodes to be simultaneously sited. For example, should we first locate one refueling node or simultaneously locate 20 nodes? Different starting numbers usually lead to different roll-out schemes.

Theoretically, without the constraint of location subset, the resulting roll-out scheme is unrealistic but provides a lower bound of *ARTT* performance. So the model is run for different starting numbers and generates the corresponding roll-out schemes with the constraint of location subset. These realistic roll-out schemes are compared with the one without the constraint of location subset with respect to *ARTT* deviation. Apparently, the smaller is the deviation, the better is the roll-out scheme.

- Multiple stations for one node

For the purpose of reducing computation time and data processing time, the total number of nodes  $|N|$  is often much smaller than the possible maximum number of stations to be considered. This means the possibility of building multiple stations around a single refueling node. Note that the distance between two adjacent nodes can be miles, so multiple stations per node here do not mean multiple stations around an intersection like 4-corner gas stations in real life, but mean multiple stations within a local area around a node. So the issue is about modelling the benefit of siting

multiple stations and trading off between building multiple stations and opening another refueling node.

The benefit of multiple stations around node  $i$  is reducing the time for those fuels aiming at the stations around node  $i$  to travel from node  $i$ 's adjacent nodes to node  $i$ . Certainly, opening a new refueling node also reduce total travel time. Thus, for any additional station to be added, the model compares these two reductions of time and choose adding a new refueling node or a multiple station, whichever brings about more reduction.

### 3. RESULTS AND DISCUSSION

#### 3.1. Station Roll-out

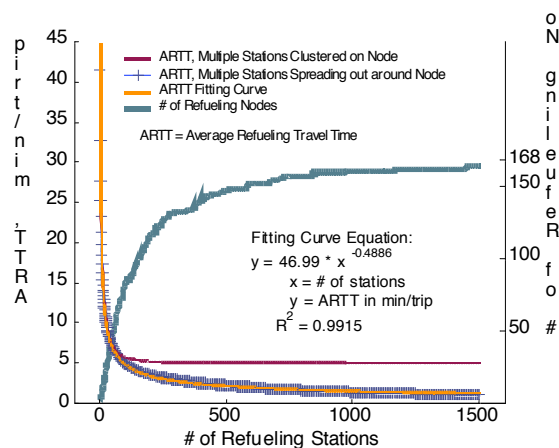
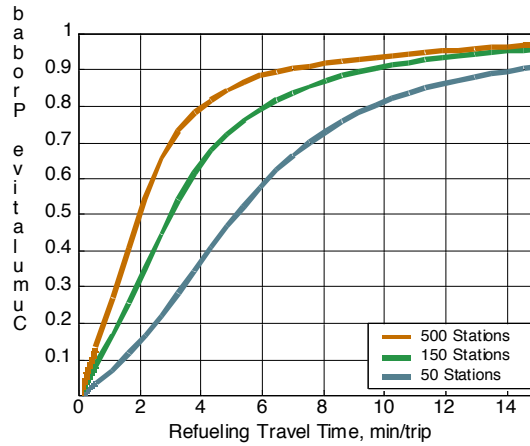


Figure 1: Average Refueling Travel Time

The number of refueling nodes increases quickly for early stations but slowly for later ones (Figure 1). This means early stations are mostly spread out over the network to reduce node-to-node travel time, and later stations are mostly spread out around existing refueling nodes to reduce node-wide travel time.

The fitting *ARTT* equation (Figure 1) indicates that travel time reduction decreases quickly with station number. The equation form is similar to the one found by Nicholas [14], indicating the robustness of equation structure. However, the fitted equation can be applied to other regions only if traffic distribution is similar.

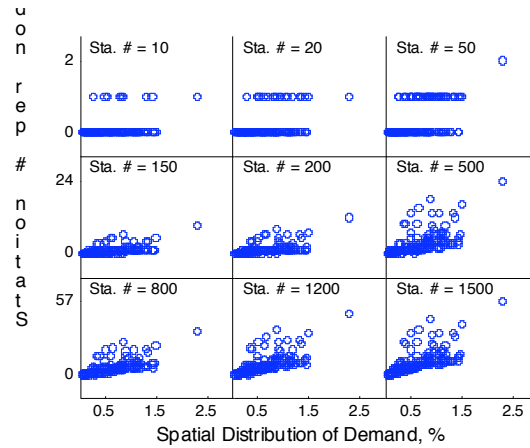




**Figure 2: Travel Time Distribution**

Although it is the optimization objective, the fuel accessibility measurement *ARTT* may not fully describe the refueling experience of a random motorist. As a supplement to *ARTT*, travel time distribution is also calculated as in Figure 2. For example with 500 stations, the expected refueling travel time for a random motorist is 2 min 16 s (Figure 1), but there is still 20% of chance that this random motorist has to travel over 4 min for refueling (Figure 2).

### 3.2. Siting Strategy for Early Stations



**Figure 3: Location Pattern**

#### **Figure 4: Station Location**

Many studies [7][17][18] agree that during the early stage, stations should be sited in close proximity to attractions, which could be high traffic volumes, high profile areas, or potential first FCV buyers. However, we find contradictory evidence in the results. The correlation between number of stations and demand share is very weak at the early stage (Figure 3). Facing the "fierce" completion among high-demand nodes for hosting one of the very few early stations, the model finds it socially better off to site a significant portion of these few stations at some intermediate nodes, instead of favoring some high-demand nodes while disregarding others. The correlation actually becomes stronger when most nodes have at least one station and the main effect of more stations is reducing node-wide average travel time (or improving local fuel accessibility). Figure 4 provides an overview of station locations for both early (50 stations) and later (500 stations) stages. These results suggest that early station location should be strategically spread out instead of close to high profile locations.

### **3.3. Station Number**

The average gasoline refueling travel time is about 1 min 50 s [14], which could vary by region. Current fuel accessibility could be achieved with much fewer stations. For the current level of 1 min 50 s, the needed number of hydrogen stations is only 708 or 18% of gas stations in the study region. This indicates a lack of location optimality of gas stations due to lack of central planning [14]. The 18% is also significantly lower than the estimate of 30% reported elsewhere [14], where locations of gas stations are the only possible locations for hydrogen stations. The implication is that, to achieve a certain level of fuel accessibility, more hydrogen stations are needed if they are restricted to gas station locations.

### **3.4. Super-large Station**

Fewer stations lead to bigger size. If 708 stations, estimated above, are to serve the whole fleet, an average size of 10,600 kg/day is required, equivalent to about 2,800 fill-ups per day. Super-large station is a possible way of utilizing economies of scale to reduce dispensing cost and provide better refueling service without sacrificing fuel accessibility. However, safety, permitting and possibly other unseen obstacles suggest further feasibility investigation.

### **3.5. Need for More Regional Studies**

The results show that a much smaller number of stations are needed, if they are optimally located. With fewer, larger stations, and therefore a more compact, lower-cost hydrogen infrastructure, it is possible to take advantage of economies of scale, leading to lower costs for dispensing, storage and distribution. This suggests that the lowest hydrogen costs might be found using specific regional data coupled with spatial optimization techniques, representing an improvement on cost estimates based on nationwide averages parameters [1][2][7][17][18] or idealized networks [19][20]. Regional studies, especially coupled with optimization techniques, could also help identify both optimistic and pessimistic regions for adopting hydrogen and could be useful to inform policy making.

## 4. CONCLUSIONS

Driven by the notion of "where you drive more is where you more likely need refueling", this paper develops a fuel-travel-back approach that only requires data of spatial distribution of *VMT*. Instead of the home or workplace, any point along the road network is a possible origin of refueling trip, with the probability quantified by *VMT* distribution. Station siting is treated as a network transportation problem. A case study for Southern California leads to the following findings.

- Average refueling travel time should be combined with travel time distribution in fully assessing refueling network performance.
- Early stations should be sited strategically and even at low-demand locations. With more stations, station location becomes more spatially correlated with demand.
- If station size constraint is relaxed, only 18% of existing stations are needed to achieve the current fuel accessibility.
- Fewer stations lead to larger station size, suggesting a need to re-examine the current speculation on designs of hydrogen station and distribution system.
- Regional studies, coupled with spatial optimization, could reduce hydrogen cost estimates and lead to discovery of optimistic and pessimistic regions for hydrogen.

## ACKNOWLEDGMENTS

The authors thank Dan Sperling, Shengyi Gao, Mike Nicholas for their help on this project. This work is funded by the Hydrogen Pathways Program at Institute of Transportation Studies, University of California, Davis.

## REFERENCES

- [1]. National Academy of Engineering (2004). The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs. Washington, D.C., the National Academies Press.
- [2]. Ogden, J. M. (1999b). "Prospects for Building a Hydrogen Energy Infrastructure." Annual Review of Energy and the Environment 24: pp.227-279.

- [3]. Sperling, D. and J. S. Cannon (2004). *The Hydrogen Energy Transition: Cutting Carbon from Transportation*, Academic Press.
- [4]. Sperling, D. and R. Kitamura (1986). "Refueling and new fuels: An exploratory analysis." *Transportation Research Part A: General* 20(1): 15-23.
- [5]. Kitamura, R. and D. Sperling (1987). "Refueling behavior of automobile drivers." *Transportation Research Part A: General* 21(3): 235-245.
- [6]. Greene, D. (1998). "Survey Evidence on the Importance of Fuel Availability to the Choice of Alternative Fuels and Vehicles." *Energy Studies Review* 8(3): 215-231.
- [7]. Melaina, M. W. (2003). "Initiating hydrogen infrastructures: preliminary analysis of a sufficient number of initial hydrogen stations in the US." *International Journal of Hydrogen Energy* 28(7): 743-755.
- [8]. Kurani, K. (1992). *New Transportation Fuels in New Zealand: Innovation, Social Choice and Utility*. Institute of Transportation Studies. Davis, University of California, Davis. PhD dissertation.
- [9]. Greene, D. (1989). "Fuel Choice for Dual-Fuel Vehicles: An Analysis of the Canadian Natural Gas Vehicles Survey." SAE Technical Paper (892067).
- [10]. Hakimi, S. L. (1964). "Optimum Locations of Switching Centers and Absolute Centers and Medians of a Graph." *Operations Research* 12(3): 450-459.
- [11]. Hooker, J. N., R. S. Garfinkel, et al. (1991). "Finite Dominating Sets for Network Location Problems." *Operations Research* 39(1): 100-118.
- [12]. Church, R. and C. ReVelle (1974). "The maximal covering location problem." *Papers in Regional Science* 32(1): 101-118.
- [13]. Bapna, R., L. S. Thakur, et al. (2002). "Infrastructure development for conversion to environmentally friendly fuel." *European Journal of Operational Research* 142(3): 480-496.
- [14]. Nicholas, M., S. Handy, et al. (2004). "Using Geographic Information Systems to Evaluate Siting and Networks of Hydrogen Stations." *Transportation Research Record* (1880): 126-134.
- [15]. Berman, O., R. C. Larson, et al. (1992). "Optimal Location of Discretionary Service Facilities." *Transportation Science* 26(3): 201-211.
- [16]. Ahuja, R. K., T. L. Magnanti, et al. (1993). *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall.
- [17]. Ogden, J. M. (1999). "Developing an infrastructure for hydrogen vehicles: a Southern California case study." *International Journal of Hydrogen Energy* 24(8): 709-730.
- [18]. Ogden, J. M., R. H. Williams, et al. (2004). "Societal lifecycle costs of cars with alternative fuels/engines." *Energy Policy* 32(1): 7-27.
- [19]. Yang, C. and J. Ogden (2007). "Determining the lowest-cost hydrogen delivery mode." *International Journal of Hydrogen Energy* 32(2): 268-286.
- [20]. H2A Analysis Group. DOE H2A Analysis. [http://www.hydrogen.energy.gov/h2a\\_analysis.html](http://www.hydrogen.energy.gov/h2a_analysis.html). Accessed in 2006.