Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/he

The deviation-flow refueling location model for optimizing a network of refueling stations

Jong-Geun Kim [a,*], *Michael Kuby* [b]

[a] *Department of Geography Education, Seoul National University, Seoul 151-742, Republic of Korea*
[b] *School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287-5302, USA*

ARTICLE INFO

Article history:
Received 7 May 2011
Received in revised form
22 August 2011
Accepted 29 August 2011
Available online 23 September 2011

Keywords:
Deviation
Optimal
Location
Station
Model
Intercepting

ABSTRACT

This research develops and applies a mixed-integer linear programming model that optimizes the locations of fueling stations considering not only the limited driving range of vehicles but also the necessary deviations that drivers are likely to make from their shortest paths in order to refuel their vehicles when the refueling station network is sparse. The Deviation-Flow Refueling Location Model (DFRLM) locates facilities to maximize the total flows refueled on deviation paths. The flow demand captured by the stations is assumed to decrease as the deviation that drivers must make increases. Test results indicate that the maximum allowable deviation and the specific deviation penalty functional form do have a measurable effect on the optimal locations of facilities and objective function values as well.

1. Introduction

Petroleum-based automobiles are the dominant mode of modern transportation systems. Concerns about Peak Oil, political instability, air pollution health hazards, and greenhouse gas (GHG) emissions of fossil fuels have challenged humankind to make a transition to a more sustainable transportation system. One of the major barriers to the success of alternative-fuel vehicles (AFVs) that use hydrogen fuel is the lack of infrastructure for producing, distributing, and delivering hydrogen [1–4]. It is clear that availability of alt-fuel stations will accelerate the market acceptance of AFVs. Based on a survey of the literature and of experts involved in hydrogen deployment, Melendez [5] identified the following as four major barriers of infrastructure development: lack of availability of hydrogen refueling stations; the high construction costs of the stations; the high costs of vehicles; and the relatively short range of vehicles between refueling. The short range of vehicles is especially relevant given the current technological state of hydrogen fuel cell vehicles. In addition, the high construction costs of hydrogen stations imply that stations in the early stages will be few and far between, and drivers will need to deviate to refuel their fuel cell vehicles (FCVs).

Efficient methods that locate hydrogen refueling stations are essential in accelerating the advent of a new energy economy. Such methods should suggest strategic station locations such that even a limited number of stations can

* *Corresponding author*. Tel.: +82 2 880 7718; fax: +82 2 882 9873.
E-mail addresses: jgkim25@snu.ac.kr, jong-geun.kim@asu.edu, jgkim25@gmail.com (J.-G. Kim), mikekuby@asu.edu (M. Kuby).

doi:10.1016/j.ijhydene.2011.08.108

achieve a satisfactory level of coverage. In addition, such methods need to be based on realistic assumptions about the characteristics of consumer demand for AFVs and drivers' refueling behavior when the stations are sparsely deployed. Kuby and Lim [6] developed the Flow Refueling Location Model (FRLM), which determines the optimal locations of refueling stations to be built in order to maximize the flows covered by a given number of facilities. The model takes into account the paths of drivers from their origins to destinations, the amount of flows on the paths, and the driving range of vehicles. In particular, the FRLM explicitly models the need for multiple stations, working in combination with each other, to refuel longer round trips. In the initial stages of the transition to alternative fuels, the lack of stations will require drivers to deviate from their regular or pre-planned routes. Berman, Bertsimas, and Larson [7] relaxed the assumption of basic flow-intercepting models—that all flows follow the shortest paths between pairs of nodes—but no one has looked at deviations for flow refueling.

The general objective of this paper is to develop a station location optimization model for decision support in deploying initial hydrogen stations. The focus of this research is to consider drivers' deviations as well as vehicle's limited driving range. The research formulates a location model that extends a flow-based location model, the FRLM to solve the problem. The model is formulated as a mixed-integer linear programming and global optimal solutions for problem instances are obtained.

This research has theoretical and societal significance. Theoretically, it provides more realistic representation of demands for station location problems. Many theories and models of facility location have been developed to serve demands that are represented as points. In addition, demand is usually assumed to be served by one facility. Recently, there is growing research interest in developing models for flow-based demands [6,8–10] or the need of multiple facilities for a full coverage of one unit of demand [6,11]. However, the provision of partial coverage of flow-based demand by a group of facilities has not been studied. There appear to be no station location models in the literature that account for deviations and vehicles' range at the same time. On a broad societal level, this work contributes to the private sector's efforts to commercialize AFVs and FCVs and the government's need to plan required subsidies. Until the AFV market becomes mature, alt-fuel stations will need to be located strategically. If public-private partnerships are involved, accessible and coordinated locations of stations will minimize the government subsidies required by fuel providers to make the delivered costs of alternative fuels competitive with gasoline by maximizing the utilization of stations constructed.

Section 1 has introduced the background of this paper and articulated the objectives and significance. Section 2 reviews related literature that estimates hydrogen demand and develops models for hydrogen refueling infrastructure. Section 3 presents the concept, assumptions, and formulation of the deviation-refueling location model along with input data generation algorithms. Section 4 describes the test network and the solution procedures, which is followed by discussion of the experimental results (Section 5). Section 6 provides conclusions and suggests future research topics.

2. Literature on modeling and analyzing hydrogen refueling infrastructure

Modeling and analyzing hydrogen refueling infrastructure entails estimation of hydrogen demands. Models for estimating hydrogen demand have incorporated a variety of assumptions and rules. In terms of their application to hydrogen infrastructure planning, the models can be grouped into five categories according to modeling method employed: logistic choice models, supply chain models, system dynamics simulation models, GIS approaches, and operation research (OR) facility location models.

In the lifecycle of an automobile, there are two interrelated cycles: the fuel cycle and the vehicle cycle [12]. To investigate the relationship between availability of alt-fuel stations and alt-fuel price, a nested multinomial logit analysis on a survey of stated preferences was used in a vehicle choice and fuel choice analysis [1]. This relationship was incorporated into a market transition model (HyTrans) that simulates the use and cost of hydrogen and FCVs [13,14]. The transition model had not accounted for spatial arrangement of H2 stations until it was integrated with the work of Melendez and Milbrandt [15] and later that of Welch [16]. The results of the integrated models were reported in Greene et al. [17], which put more emphasis on the phased rollout of refueling stations in regions of high potential demand. A similar model and projection was also suggested in Keles et al. [18] for estimating FCV demand in Germany. With regard to ways of estimating demand, all variants of the model estimate hydrogen demand from the given number of FCVs by multiplying it by a constant—average hydrogen consumption per vehicle—which was derived from the Hydrogen Analysis Project (H2A) [19].

In examining the hydrogen supply chain from production center to refueling stations, Ogden [19] built a database on the costs of delivery system components and applied it on an idealized city model (ICM) with with regularly sited stations of the same size, which was revised in Yang and Ogden [20] to model the lowest-cost hydrogen delivery option from a large central production plant to vehicles.

Welch [16,21,22] analyzed consumer sensitivity to alt-fuel station coverage using a discrete choice model. The choice model was implemented in a system dynamics simulation model (HyDIVE) to estimate required vehicle price, vehicle makes, and fuel cost to meet DOE's [17] FCV market sales target. In the model, demand is proportional to population and a log-normal driver trip frequency distribution is assumed. In the sensitivity analysis of the choice model, station convenience attributes are measured in terms of mid- or long-distance trips (20–150, +150 mile from home) rather than percentage of existing stations.

GIS-based approaches for estimating hydrogen demands generally used two types of data: existing gasoline stations and demographics. Based on data of existing conventional stations, a series of studies [23–27] have focused on various aspects of size, number, and spatial configuration of refueling stations needed for AFVs. On the other hand, without detailed data of conventional fuel demands, others used more available demographic data with GIS in estimating hydrogen demand. Ni et al. [28] used GIS to determine hydrogen center demands given a penetration rate in the vehicle stock in a region. Melendez and Milbrandt [29] developed a plan for a national hydrogen network to make interstate trips possible by regularly placing stations along highways in areas of high potential hydrogen demand. In a series of analyses, Melendez and Milbrandt [15] identified the areas of highest FCV demand, and then assessed how many stations would be needed to fuel these vehicles and where they might realistically be located.

Many approaches reviewed above do not provide an optimal solution that satisfies a given set of conditions by examining all possible combinations of locations. Optimization-based approaches for locating refueling stations are divided here into two groups depending on the geometric representation of demand: models for point-based demands and flow-based demands. Most of the classical facility location models [30–36] involve providing service to point-based demands. Bapna et al. [37] used a multi-objective model to locate unleaded gasoline stations in India. One objective minimizes the sum of travelers' costs and station investment costs, while the second maximizes the population on enabled links. Bersani et al. [38] took into account competition by incorporating the Huff [39] model in a multi-objective model. Many median-based models have been used for siting refueling stations. Goodchild and Noronha [40] developed a model to decide which gasoline stations of a firm to keep open or close to maximize market share. They recognized that refueling trips are composed of a mix of traffic-originated demand and population-originated demand. Their model was in essence a p-median problem that minimizes the two different distributions of demand. Chan et al. [41] reported on government use of the p-median model to locate gasoline stations in Singapore. Lin et al.[42] proposed the fuel-travel-back model, which uses link traffic as the weight and locates stations to minimize the sum of average weighted distance. The model is structurally identical to the p-median model with link traffic as the weight at each road intersection. Analysis by Nicholas et al. [43] using a p-median approach determined the number of stations by the average driving time per refueling trip. It is important to note that, if population is used as a weight in a model based on the p-median, the implicit assumption is that drivers travel from home to a station and back for refueling, which is arguably not a typical refueling behavior. Even though, by using link traffic as the weight, the model may partially account for drivers' more typical behavior of refueling when needed while driving, link traffic might be doubly counted by more than one station along the path, as pointed out by Hodgson [8]. Such double counting of traffic volumes could lead to duplicative siting and cannibalization of a station's demand by other stations. In addition, given the limited range of AFVs, p-median-based models may not site stations to enable a long inter-regional (e.g., LA to San Francisco) trip. This is one benefit of using flow-based demand in models for locating stations of range-limited AFVs.

Recently there has been increasing research interest in flow-based demand that is expressed by flows traveling on paths between origin–destination (O–D) pairs in a traffic network. The flow-intercepting location model (FILM) sites facilities within a transportation network and explicitly considers the flow over the network arcs. Refueling stations [6], convenience stores, and automated teller machines, vehicle inspection stations [9,44], and billboards [45] are examples of flow-dependent facilities. Hodgson [8] and later (independently) Berman, Larson, and Fouska [46] designed the Flow Capturing (Intercepting) Location Model (FCLM, FILM) to locate these kinds of flow-dependent facilities. The objective of this model is to locate the facilities so as to maximize the total flow of customers that are "intercepted" during their travel. The basic model was extended to provide full coverage [47], consider relative location of facilities along the path [48], or account for congestion or probabilistic flows [49]. Berman et al. [7] relaxed a key assumption required by the basic FCLM: that customers can make no deviation, no matter how small, from their pre-planned paths to visit the facilities.

Kuby and Lim [6] extended the FCLM to locate a given number of facilities to maximize the number of flows they can refuel. The new model (FRLM: Flow Refueling Location Model) is intended to deal with location of refueling stations for range-limited vehicles, with vehicle range being the key element. A limited driving range means that one facility anywhere on the path cannot necessarily succeed in refueling a trip on a given shortest path—a combination of facilities may be needed. Whereas the FCLM counts a flow as captured if a facility is located anywhere along the path of the flow because one stop will satisfy consumers' need, the FRLM regards a flow as refueled only when a satisfactory number of facilities (stations) are spaced properly along the path because consumers on the path may need multiple stops. To improve the FRLM's solution quality for a given network, Kuby and Lim [50] proposed methods to add candidate locations along arcs. Upchurch et al. [51] extended the FRLM to consider capacity of facilities. Given that solving a FRLM problem instance to optimality is computationally challenging, heuristic algorithms [52] and a new formulation of the model [53] were also proposed. The FRLM was used to provide strategic station locations for the state of Florida at two different scales of analysis: metropolitan Orlando and statewide [54].

Flow-based models can be improved to become even more realistic and more widely applicable by allowing drivers to deviate from their shortest paths. It is expected that this relaxation will increase each station's utilization level, and in turn it will result in less investment costs and a more successful transition to a new transportation system.

3. Deviation-flow refueling location model

The key aspect of this paper is to relax the assumption of the flow refueling location model that customers do not deviate from their pre-planned trips to refuel their vehicles. In the new model, locations *somewhat close* to pre-planned paths may be able to serve them. This is believed to better reflect drivers' behavior when the refueling network is sparse and the required deviation to the facility is acceptable. We assume that the number of customers willing to deviate to a station off their pre-planned path decreases as a function of required deviation distance.[1] The objective of the new model, which is called the Deviation-Flow Refueling Location Model (DFRLM), is to locate p facilities to maximize the expected number of potential trips using range-limited vehicles that can be refueled at the facilities.

An important input for the new model is *all* deviation paths between O–D pairs. Given that the problem of finding all possible paths between an O–D pair in a real-world network is hard to solve, the generation of feasible deviation paths requires additional constraints that reflect more realistic assumptions on drivers' refueling behavior in a sparse refueling network. In addition, without available empirical data as to how many AFV drivers would deviate from their pre-determined paths and how far they would do so for refueling purposes, an explorative model is required. With such a model, planners could perform experiments with different upper bounds of deviation distance or apply different functional forms to fit the decreasing willingness of drivers to deviate from their paths as deviations get larger.

Finally, all of these model extensions need to consider vehicle range at the same time. Kuby and Lim [6] pointed out that enforcing a driving range restriction requires that the driver be able to not only make it from the origin to the destination without running out of fuel, but also be able to return.

3.1. An example network

We illustrate the difference among the FCLM with deviation, the FRLM, and the suggested DFRLM model with a simple network with 5 nodes and 2 paths (Fig. 1). For simplicity, all path flow will be assumed to be 100 as in Table 1. Suppose that the driving range of an AFV is 100 unit distance and we want to locate two stations ($p = 2$) with one station already having been located at node A. The objective is to maximize the covered flows by the two stations. If the FCLM with deviation is applied in this situation, the additional station is not even needed because the station at node A will be able to capture all the flows of path 1 (A–B) and 2 (A–E). However, when the basic FRLM (with no deviations allowed) is used to solve the problem, path 1 cannot be served by node A alone because of the vehicle's range limit but it can be captured only by adding an additional station at node B. Likewise, in order to capture the flows on path 2, a new station would need to be built at node E. Note that it is impossible to provide service for both paths by building only one more station at either node and therefore, solution $\{A, B\}$ is equally good as $\{A, E\}$. However, when deviations from the shortest path are allowed for the FRLM, choosing either C or D can provide service for both paths. Node C can refuel path 1 with deviation distance 40 ($\overline{ABCBA} - \overline{ABA}$) and path 2 with deviation distance of 120 ($\overline{ABCECBA} - \overline{ADE}$). The customer flows refueled by node C, however, should be less than the sum of each flow (200) because the number of visiting customers should decrease as deviation distance increases. Locating a station at node D can refuel path 1 with deviation distance of 20 ($\overline{ADBDA} - \overline{ABA}$)

[1] Deviation distance is directly calculated by subtracting shortest path distance from deviation path distance.

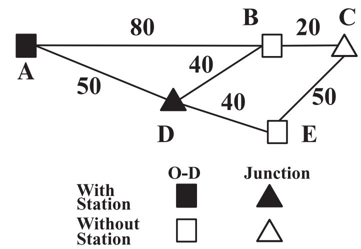

Fig. 1 – A 5-node network.

while serving paths 2 without any deviations. When evaluating C and D, locating the station at D is likely to be superior to locating at C because of its higher objective value (this of course depends on the deviation penalty function). The solution for deviation case $\{A, D\}$ should provide higher objective than the solution for basic FRLM, $\{A, B\}$ or $\{A, E\}$, does because the former accounts for the additional customer flows that deviated from their shortest paths. This example shows that allowing deviations in modeling flow refueling will increase the utility of stations by serving more customers.

3.2. Assumptions

3.2.1. Shortest deviation path

An assumption related to the deviation distance calculation is that drivers take the shortest or least cost deviation path to their required refueling stations and then to their final destination. That is, we assume they have perfect information and choose the best way to deviate from their shortest path to the off-path station(s) and then to their final destination. Such behavior is expected to be more common in the future due to the increasing availability of onboard vehicle navigation systems. In sum, we assume that drivers who need to take a deviation path for refueling may have to visit multiple stations and the sequence of them has to be optimized so that the total deviation path distance is minimized.

3.2.2. Flow volume decay with the increase of deviation

Another assumption relates to modeling of penalties on deviations. This has been handled in a few different ways in FCLM-deviation models [7,55]. One possible assumption is that a given Δ unit of distance away from a node in the pre-planned trip path is regarded as not affecting the flow volumes. If the driver then returns to their shortest path, this assumption implies that all customers are indifferent to up to 2Δ unit deviation from their original paths. Alternatively one may choose not to impose an upper limit on deviation

Table 1 – O–D flow paths.

Origin–destination pair q	Flow	Shortest path by nodes	Shortest round trip distance (time)
1 (A–B)	100	A–B	160
2 (A–E)	100	A–D–E	180

distance no matter how long it may be, as in Hodgson [55]. Another possible assumption is that the fraction of customers deviating to a facility can be specified as a decreasing function. Here we assume the latter, namely that capturing of flows by stations off the shortest path declines from 100% to 0% according to some exogenously given deviation penalty function.

This paper does not assume a specific curve shape of the decreasing function to describe the fraction of flows on deviation paths, but loosely defines it as a decreasing function of deviation distance. The reliability of a penalty function is highly dependent on the specific situations to which the model will be applied. Previous research employed exponential decay [7,48]. In the transportation literature, inverse distance functions have been widely accepted. The exact shape and steepness may need to be calibrated to empirical survey data, which is beyond the scope of this paper. Note that depending on the specification, the extent to which customers show indifference to the required deviation could also be modeled by allowing users to specify parameters for five different types of distance decay functions: no distance decay, linear, exponential, sigmoid, and inverse distance. Design and implementation details are described in Section 3.4.

In addition, we assume that there is an upper limit of deviation distance that drivers can tolerate, even though the path might be refuelable.[2] It may be a psychological limit that is dependent on consumers' perception on deviation.

Both distance decay function and upper limit of deviation distance can be specified in relative or absolute terms. In the case of a penalty function, the spatial extent of distance decay is referred to here as bandwidth. Large bandwidth implies more gradual distance decay, where a small bandwidth results in a rapid decrease. The bandwidth can be a fixed distance or can be relatively derived from each reference path distance (i.e. some percentage of the shortest path distance). Likewise, the upper limit of deviation distance can be a fixed value or it can be specified as percentage of shortest path distance.

3.2.3. *Common assumptions of flow refueling location model*

The DFRLM is an extension of the FRLM, and therefore it follows assumptions of its precedent model, which are presented in Kuby and Lim [6]. The following is a summary of them.

- Flow refueling location models are formulated to locate facilities that make round trips feasible.
- The starting level of a vehicle's fuel is half the fuel tank, which ensures that the round trip will be feasible and repeatable.
- Fuel consumption is strictly a function of distance.
- Facility location is limited to network nodes. It will be, however, possible to extend DFRLM to add candidate locations at locations along the links as the extensions of FRLM do.

[2] While an upper limit is not necessary when using the linear decay function that hits zero at a finite deviation length, they are necessary for asymptotic decay functions.

3.3. *Formulation of the deviation-flow refueling location model*

This section presents a mixed-integer linear programming formulation of the deviation-flow refueling location model (DFRLM). The model relaxes the FRLM [6] such that it allows driver's deviations to the facilities that are near customers' pre-planned shortest paths as considered in Berman et al. [7]. A new subscript r is introduced to represent deviations.

Formulation of the DFRLM

$$\text{Maximize} \quad \sum_q \sum_r f_q g_{qr} y_{qr} \tag{1}$$

Subject to

$$\sum_{r \in R_q} y_{qr} \le 1, \quad \forall q \in Q \tag{2}$$

$$\sum_{h \in H_{qr}} v_h \ge y_{qr}, \quad \forall r \in R_q, \ q \in Q \tag{3}$$

$$x_k \ge v_h, \quad \forall h \in H, \ k \in K_h \tag{4}$$

$$\sum_{k \in K} x_k = p \tag{5}$$

$$x_k, v_h, y_{qr} \in \{0,1\}, \quad \forall k \in K, \ h \in H, \ q \in Q, \ r \in R_q \tag{6}$$

where:

Indices
q = a particular O–D pair (the shortest path for each pair)
r = index of deviation paths
k = a potential facility location
h = index of combinations of facilities

Sets
Q = set of all O–D pairs
R = set of all deviations
R_q = set of deviation paths r for O–D pair q
K = set of all potential facility locations
K_h = set of facilities k that are in combination h
H = set of all potential facility combinations
H_{qr} = set of facility combinations h that can refuel deviation path r that is originated from O–D pair q

Parameters
p = the number of facilities to be located
f_q = flow between O–D pair q
g_{qr} = fraction of normal path q customers who would be willing to take deviation path r (that is, the penalty function value for deviation r)

Decision Variables
x_k = 1 if there is a facility at location k, 0 if not
y_{qr} = 1 if path r is the highest-volume path for O–D pair q that can be refueled, 0 otherwise
v_h = 1 if all facilities in combination h are open, 0 otherwise

The objective function (1) maximizes the total flow that can be refueled. Constraints (2) limit the contribution to the objective function to at most one deviation path r. The set of paths r for O–D pair q includes q itself, as q is the shortest path from O–D pair q with deviation distance = 0. The number of possible r can be very large, and thus without this constraints, double counting of flows on deviation paths might also be possible. Constraints (3) ensure that for any deviation path r to be refueled, at least one valid combination h has to be open. The v_h variables serve, in a sense, as intermediate variables linking one or more x_k variables, representing the decisions of where to build stations, to a single y_q variable representing the decision of which paths to cover. Determination of the eligible combinations of facilities for each shortest or deviating path is exogenous in that it is generated outside the model and depends on the network structure and the given vehicle range. An algorithm to generate the combination h for each path q and other considerations such as obtaining a tighter set H by removing supersets are discussed by Kuby and Lim [6] and it also can be used for deviation path r. Constraints (4) ensure all the facilities in combination h are open before v_h becomes one. Constraints (5) specify the number of facilities to open, and (6) are the integrality constraints for the variables.

3.4. *Algorithms to generate input data for DFRLM*

Three algorithms were implemented to prepare input data for the MILP formulation of DFRLM (Fig. 2). One algorithm generates deviation paths for each O–D pair given the upper limit of deviation distance. Another algorithm computes the fraction of flows on deviation paths. This algorithm reads in the user's input parameters to formulate a distance decay function. The last algorithm determines whether a particular combination of facilities can refuel a deviation path. The first two algorithms are new to the FRLM literature, while the third algorithm is from Kuby and Lim [6] except that it runs on all deviation paths rather than on shortest paths. The three algorithms are implemented in an application using the C# programming language.

3.4.1. *Generating deviation paths: modified* k*-shortest path algorithm*

This algorithm generates the set of deviation paths R and R_q. It first reads in an upper limit of deviation distance, which could be in relative term or absolute term. The next step is to run Hoffman and Pavely's [56] k-shortest paths (KSP) algorithm with $k = 1$, and then the algorithm evaluates if its deviation distance is shorter than the upper limit. If so, the algorithm increases k by 1 and keeps generating alternative paths by running the KSP algorithm until the deviation distance reaches the upper limit. Note, however, that if the upper limit equals 0, no deviation is allowed, and the KSP generates only shortest paths, which in effect converts the DFRLM into the FRLM. It is, however, possible that the KSP may generate multiple paths with the same distance, in which case one may consist of more edges with shorter distances while others are made up of fewer edges with longer distances. Even in such cases, only one deviation path contributes to the objective value because constraints (2) in the MILP of the DFRLM ensure the inclusion of at most one deviation path per an O–D pair. The algorithm stops when all O–D pairs are evaluated.

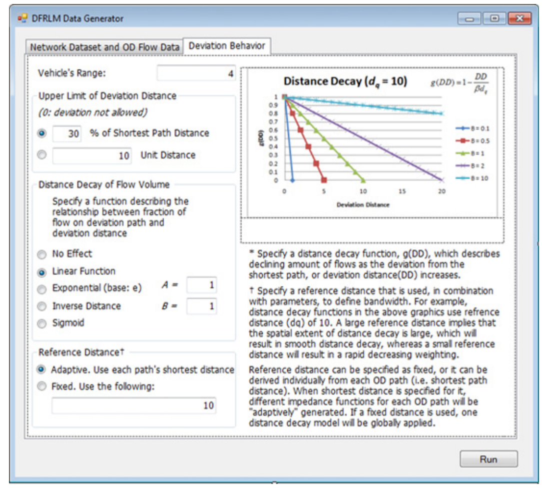

Fig. 2 – Input data generation for DFRLM.

Two minor simplifying assumptions are made in applying the KSP. Although these simplifications imply that some unrealistic paths are generated and some theoretically possible round trips are not generated, we believe the effect on the model in practice is negligible. One modeling issue is whether or not deviation paths from an origin to a destination should contain loops. In many implementations of KSP algorithms, loops are not allowed. Yen [57], Dreyfus [58], Shier [59], and Eppstein [60] provide more complete reviews on KSP algorithms. If we consider the refueling behavior when stations are scarce, however, the only feasible path may require drivers to take a short deviation containing a loop in order to reach a fuel station. For example, if there is a station at node C (Fig. 1) that is close to the destination of path 1, drivers may take a path A–B–C–B to reach the station at C and then loop back to the destination B. Therefore, it is important that the KSP used to generate deviation paths for the DFRLM consider paths with loops. When loops are allowed, however, a KSP algorithm might generate a path with multiple loops. For example, in Fig. 1, paths A–B–C–B and A–B–C–B–C–B could both be generated as different deviation paths. Given that drivers typically would take only the single loop necessary to refuel, removing multiple loops in the KSP algorithm would be more preferable to reduce the size of problem. In most cases, however, allowing multiple loops in the KSP will not lead to erroneous results, because the set of facilities that can refuel the former is also able to refuel the latter path, and *any* downward-sloping penalty function g_{qr} used in the objective function coefficients will favor the shorter deviation path and ensure that the longer deviation path is never optimal.

The second necessary simplification is related to the FRLM's round trip assumption. Most KSP algorithms generate one-way routes, but do not usually find a round trip path from an origin to a destination and back. However, it is possible that a driver's egress path might need to be different from their ingress path. For example, with a driving range of 100, the shortest refuelable ingress path q_{in} in (Fig. 3) is O–A–R_1–D, if a driver deviates from the shortest path to refuel at R_1. In this case—with this network and this driving range—q_{in} is feasible

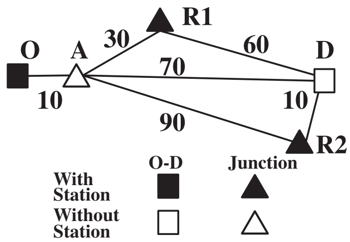

Fig. 3 – Non-symmetric round-trip path.

but the vehicle arrives at *D* with only enough fuel for 40 distance units, making it infeasible to visit the nodes of q_{in} in the reverse order. However, it *is* feasible to return to O via the egress path q_{out} consisting of $D–R_2–A–O$, refueling at R_2, and again at O at the end of the round trip, which makes any subsequent trip possible. A standard traveling salesman problem (TSP) algorithm cannot be used to generate these types of round trips because the nodes (stations) to be visited in a deviation-refueling path are not pre-determined as in TSP. Another option would be to generate all possible combinations of ingress and egress paths, but this would greatly multiply the number of deviation path variables to consider in order to accommodate a rare circumstance requiring creative driver navigation. This paper, therefore, does not allow non-symmetric deviation round-trip paths and drivers are assumed to visit the nodes of their ingress paths in the reverse order when returning to their origins.

3.4.2. *Computing the fraction of flow volume on deviation paths*

This algorithm computes the fraction of path *q* customers who take deviation path *r*, which is denoted as g_{qr} in the MILP formulation of the DFRLM. Let *DD*, d_q, and $g(DD)$ be the deviation distance, reference distance, and the fraction of flows specified by function *g* and input *DD*. Users can choose a distance decay function type from five available types: no decay, linear, exponential, inverse distance, and sigmoid. Each function can be expressed as below:

$$g(DD)_{\text{No Decay}} = 1 \tag{7}$$

$$g(DD)_{\text{linear}} = 1 - \frac{DD}{\beta d_q} \tag{8}$$

$$g(DD)_{\text{exponential}} = 1 - \left(\alpha e^{\beta(DD - d_q)}\right) \tag{9}$$

$$g(DD)_{\text{inverse distance}} = \left(\alpha e^{-\beta \cdot DD}\right) \tag{10}$$

$$g(DD)_{\text{sigmoid}} = \frac{1}{1 + \alpha e^{(\beta \cdot DD) - d_q}} \tag{11}$$

The fraction of flow volume on a deviation path with deviation distance of *DD* is expressed as a function of reference distance d_q and deviation distance *DD*. By specifying parameters (α or β), the shape of the function is determined. Fig. 4 shows four classes of graphs depicting deviation function shapes with reference distance of 10 and different parameters. These illustrative graphics are provided to aid users' specification of distance decay function. Note also that by substituting other constants for *e* in (11), the steepness of the transition from near 1 to near 0 can be controlled.

The reference distance, in combination with the other specified parameters, defines the bandwidth. For example, the distance decay functions in the example use reference distance d_q of 10. A large reference distance implies that the spatial extent of distance decay is large, which will result in smooth distance decay, whereas a small reference distance will result in a rapidly decreasing weighting. It can be specified as fixed, or it can be derived individually from each O–D path

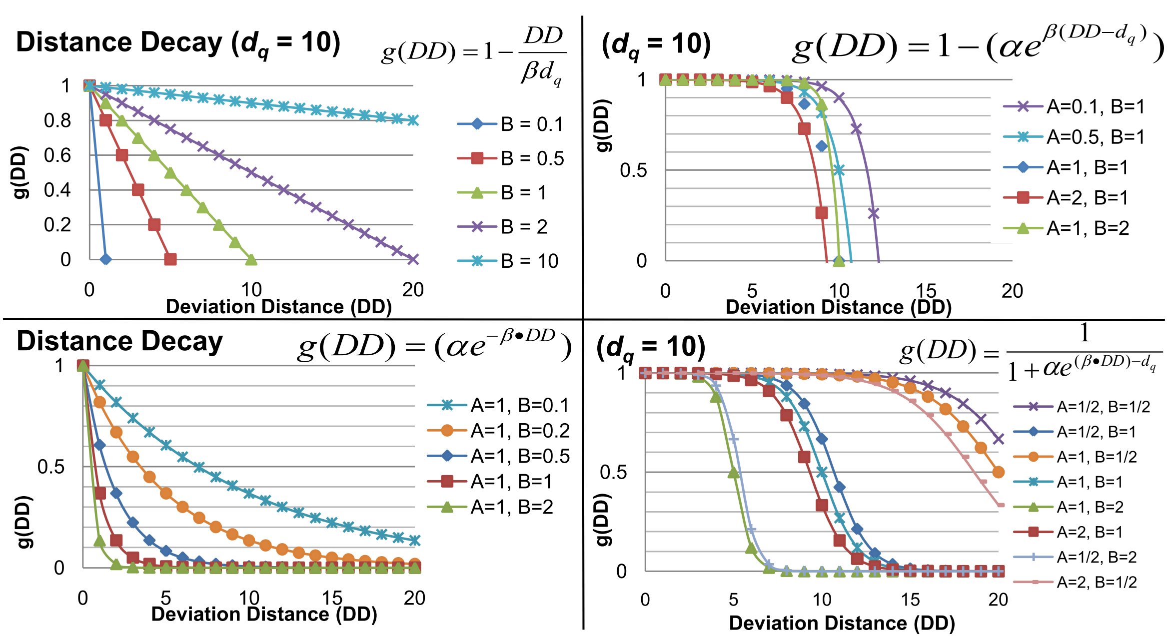

Fig. 4 – Example of distance decay functions.

(i.e., shortest path distance). When the shortest path distance is specified, different impedance functions for each O–D path will be "adaptively" generated. If a fixed distance is used, on the other hand, one distance decay model will be globally applied.

3.4.3. *Evaluating feasibility of deviation path*

The MILP formulation of the DFRLM requires generation of the valid set of facilities h for each deviation path r. The formal algorithm for this task is presented in Kuby and Lim [6] in detail. The difference is that for the DFRLM, the set H is generated over all the deviation paths rather than all the shortest paths. The algorithm reads in the vehicle range and the sequence of nodes on a deviation path for an O–D. The remaining range in a vehicle is set to half the vehicle range unless there is a station at the starting node. It moves to the next node in the sequence while subtracting the link distance traveled from the vehicle's remaining range. The remaining range is assumed to be full when the visiting node already has a facility. If the round trip taking the deviation path of the O–D pair successfully ends without running out of fuel, the deviation path is considered feasible with the given set of facilities. The algorithm runs for all deviation paths and outputs the list of r and its corresponding h. Because this algorithm assumes that drivers refuel whenever there is a station, it works regardless of the number of loops in the deviation paths as long as the path is feasible.

4. Test network and solution procedure

The DFRLM was tested on a network with 25 nodes, 43 edges, and 300 O–D pairs that was used by Berman and Simchi-Levi [61], Hodgson [8], and Kuby and Lim [6] (Fig. 5). The flow volumes were estimated using a gravity model, and assigned

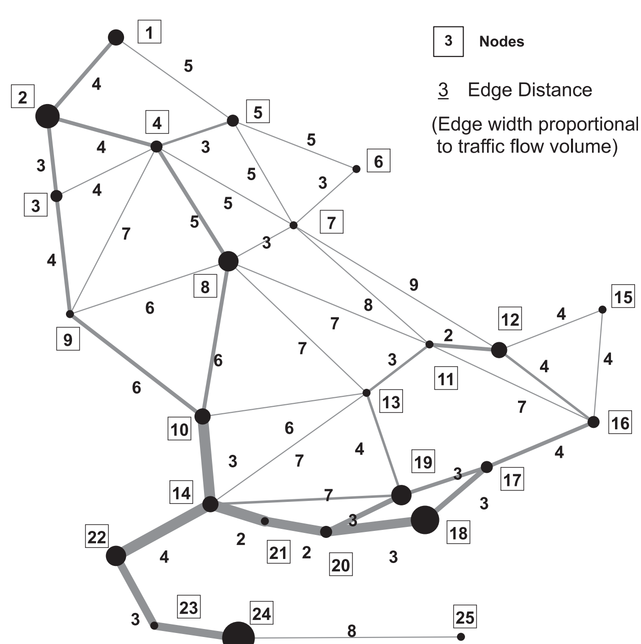

Fig. 5 – The test network with 25 nodes.

to their shortest paths. Each node is both an O–D node and a candidate site.

The MILP of the DFRLM was built in and solved optimally by using Xpress 7.0 software, and the input for the model was generated by an application coded in C# language, described above. For the upper limit of deviation distances,[3] three values were used: 0%, 10%, and 50% of shortest path distance. To account for the fraction of flows on deviation paths, $g(DD)_{\text{No Decay}}$ and $g(DD)_{\text{linear}}$ were used. In the linear function, bandwidth is the shortest distance of each O–D pair, and therefore, the fraction of flows becomes 0% when deviation distance equals shortest path distance ($DD = d_q$). Three different vehicle ranges (4, 8, and 12) were used. All the problem instances were solved by a computer with four 2.4 GHz cores and 4 GB memory.

Problem size is highly dependent on the number of deviation paths $|R|$, which is mainly determined by the maximum deviation distance allowed ($DD_{\max}$). It is also dependent on vehicle range, which will determine the size of the set of potential facility combinations $|H|$. For instance, when vehicle range is 4 and no deviation is allowed, there are 395 deviation-flow coverage variables Y_r; 109 facility combination variables V_h; 67 contributing deviation path choice constraints (2); 119 refueling constraints (3); and 396 combination constraints (4). On the other hand, if the $DD_{\max}$ is allowed up to 50% of shortest path distance and vehicle range is set 12, the number of variables and constraints increase substantially: 15,344 Y_r; 13,883 V_h; and 90,992 total constraints.

5. Numerical experiments

5.1. *Computation time*

Computation times for each test problem are summarized in Table 2. In general, as the maximum deviation distance gets higher, more time is required to generate deviation paths. The branch-and-bound solution time is shorter when linear distance decay function was used. It is also faster when the maximum deviation distance is lower, or the vehicle's range is shorter.

In an attempt to solve the problems faster, only X_k is forced to be binary and other decision variables (Y_{qr}, V_h) are relaxed to be continuous between 0 and 1. Because the MILP formulation of DFRLM is unimodular (as is the FRLM [6]), and satisfies ReVelle's [62] definition of "integer friendly," this relaxation will not affect the global optimality but will result in a reduced number of binary variables, which in turn enables the model to solve faster than the all-integer version of the problem. We randomly chose three problems to observe the computation time reduction entailed by the

[3] A cautionary choice of upper limit deviation distance is required. For example, given that the farthest O–D pair is 1–25, whose shortest path length is 38, 100% of deviation means drivers are willing to deviate up to 38 distance units on this O–D pair. With this deviation distance, the number of possible deviation paths is enormous, and thus the upper limit should not be very high. In fact, Xpress returned a "not enough memory" error for the instance of DFRLM with range 8 and upper limit of 100% of the shortest path.

Table 2 – Computation time for test problems for p = 1–25 (CPU seconds[a])

Vehicle range	Procedure	No deviation		Distance decay function: linear		Distance decay function: no decay	
		One SP	Multi- SP	DD_{max}: 10% of SP	DD_{max}: 50% of SP	DD_{max}: 10% of SP	DD_{max}: 50% of SP
4	B–B	0.921	1.39	1.72	11.95	1.594	13.567
	Data generation	N/A	0.47	0.54	59.71	0.533	59.302
	Total	0.92+	1.85	2.25	71.66	2.127	72.869
8	B–B	5.6	6.88	14.60	1279.20	12.551	1441.618
	Data generation	N/A	0.53	0.66	159.70	0.647	163.91
	Total	5.60+	7.41	15.26	1438.90	13.198	1605.528
12	B–B	8.72	11.341	23.22	2511.14	22.424	2166.406
	Data generation	N/A	0.623	0.83	493.57	0.82	494.132
	Total	8.72+	11.964	24.05	3004.70	23.244	2660.538

B–B = branch and bound and SP = shortest path.
a Based on mixed-integer variable formulation of DFRLM.

relaxation. The time reduction ranged from 2.6% to 72% of the computation times of the all-integer versions with more reductions for larger problems.[4]

Given that the test network has 25 nodes, the ability to apply the MILP of the DFRLM to a real-world network may be limited to very small networks, which highlights the need for efficient heuristic methods.

5.2. *Effects of vehicle range and lack of convexity*

As is observed in [6], full refueling of all O–D pairs is not possible even with facilities open for all candidate node locations when the vehicle range is too small (range = 4 or 8). This is because both DFRLM and the original FRLM are restricted to consider nodal points as candidate sites. On paths containing a link that is longer than the vehicle's range, no flows across that link can be refueled even if stations are located at both nodes. This still applies to these deviation cases. When the range is 12, however, the total flows can be refueled with 15 stations if all the drivers are assumed to willingly deviate up to 50% of their shortest path distances.

The tradeoff curves in Figs. 6–10 are not convex even though they are non-decreasing. As stations are added to the solution, the marginal coverage enabled by the added station does not necessarily decrease from that of the previously added station because there are cases where the addition of one station enables a combination of facilities to refuel some flows that would not be feasible otherwise. This is also a property common to the nodes-only version of the FRLM and its extensions [6,50].

5.3. *Effects of deviation distance and distance decay function*

Tables 3 and 4 summarize the test results. By looking at how much the objective value improves compared with the base value obtained from the no-deviation case, it is clear that as the deviation distance increases, the same number of stations covers more flows (Fig. 6). The percentage improvement ranges from 0.07% up to 14.79%. The most improvement was obtained when the stations' coverage is between 60% and 80%. We can interpret this range as providing more alternative paths with combinations of built stations. When stations are scarce, refueling trips would require a long deviation, which however may not be feasible because of unavailability of stations. As the stations are more available but not so many to cover the most of the flows on the shortest paths, drivers will still need to deviate to refuel and there could be feasible combinations of open stations that enable the O–D flows. These deviation flows are not necessary when the refueling network becomes mature and most flows can be refueled on their shortest paths.

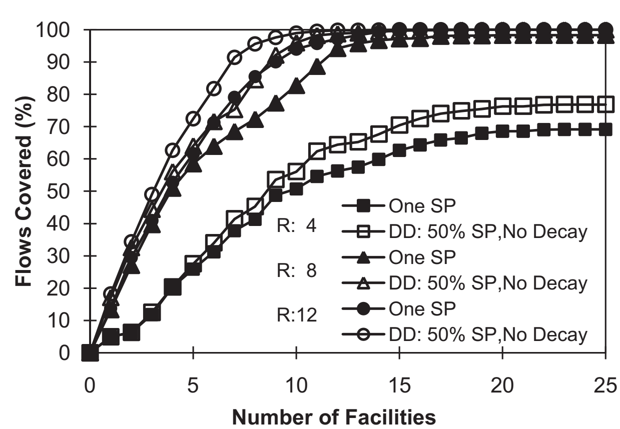

Fig. 6 – Tradeoff curves for contrasting sensitive to deviation.

[4] We confirmed that relaxing the 0–1 Y_{qr} and v_h variables always yielded the same objective function value as when branch-and-bound was used to force them to 0 or 1. In most problems tested, the relaxed Y_{qr} and v_h variables solved to 0s or 1s anyway because the objective function will try to maximize the feasible Y_{qr} with the smallest deviation. Occasionally, however, they yielded fractional solutions in the case of ties between two or more deviation paths with the exact same deviation. Ties are not uncommon in our 25-node network with whole-number arc lengths, but are highly unlikely to happen with real-world distances. In cases where it does occur, there exist several all-integer alternative optima, which can be found by rounding any fractional v_h to 1 and randomly choosing one of the fractional Y_{qr} for a given q to round up to 1 while rounding the others to zero.

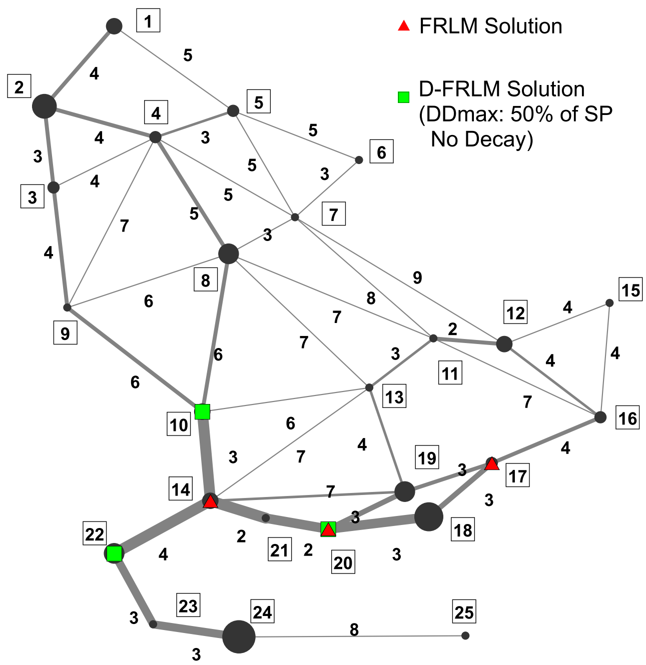

Fig. 7 – Optimal solutions of DFRLM and FRLM for $p = 3$, range = 12.

Table 4 contrasts different assumptions of drivers' sensitivity to their deviation distance. When drivers are indifferent to deviation distance up to 50% of SP and the range is 8, full coverage of the flows is possible with 19 stations, whereas the sensitive flows cannot be fully refueled even with 25 stations when restricted to their shortest paths (labeled One SP in Table 4). Similarly, 15 stations can cover the full demand with a range of 12 while the full coverage of these sensitive flows requires 17 stations. This complementary information about sensitivity to deviation distance will be useful for the alt-fuel infrastructure service providers to strategically determine the level of coverage and the number of stations to construct. Table 4 and Fig. 6 reveal an interesting point that a smaller number of stations can refuel more indifferent flows of range 8 than sensitive flows of range 12 in as many as 12 out of 19 p values.

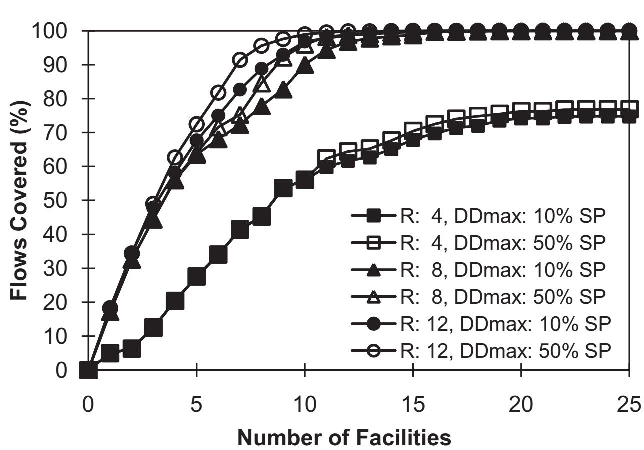

Fig. 8 – Tradeoff curves for different deviation distances.

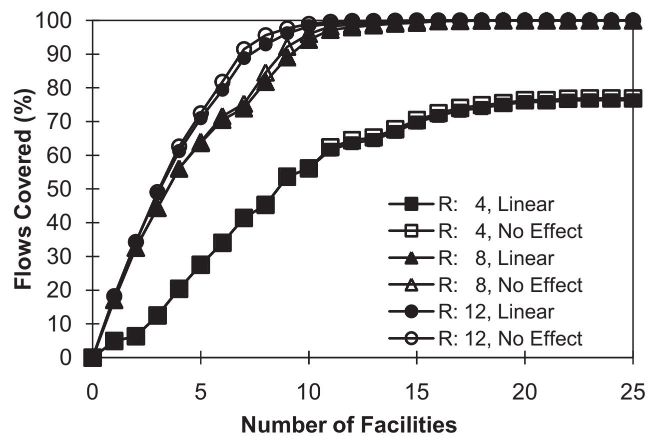

Fig. 9 – Tradeoff curves of different distance decay functions.

Fig. 7 shows a typical difference in the solutions of DFRLM and FRLM; where both located three facilities for vehicles of range 12. The DFRLM solution refuels 8.12% more O–D pairs by serving the flows on deviation paths (Table 3). When the stations are at nodes 10, 20, and 22, the O–D pair of 8–17 can be refueled by the open facilities if drivers are taking a deviation path: 8–10–14–21–20–18(19)–17, the distance of which is 19. Because the shortest path should visit 8–13–19–17 and its length is 14, the deviation distance (19–14 = 5) is less than 50% of the shortest path length. Therefore, DFRLM considers that all the flows of the O–D pair will be covered by the solution. There are more such O–D pairs that the FRLM solution cannot take into account. The O–D pair 13–14 is another example.

Figs. 8 and 9 show the effect of the allowed maximum deviation distances and distance decay functions. Obviously, deviation distance has direct impact on the coverage level (Fig. 8). Increase of deviation distance results in higher objective values. In addition, deviation flows estimated by applying a $g(DD)_{\mathrm{NoDecay}}$ yielded a higher objective value than by $g(DD)_{\mathrm{linear}}$ (Fig. 9).

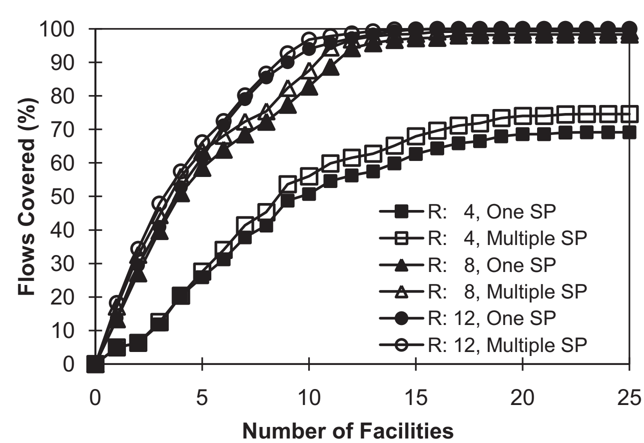

Fig. 10 – Tradeoff curves for the same problems with different formulations (FRLM: one SP/DFRLM: multiple SP).

Table 3 – Optimal coverage gain with different ranges (4, 8, and 12) and deviation functions.

p	Range of 4					Range of 8					Range of 12				
	% of Flows refueled with no devia-tion[a]	Distance decay function: linear		Distance decay function: no decay		% of Flows refueled with no devia-tion[a]	Distance decay function: linear		Distance decay function: no decay		% of Flows refueled with no devia-tion[a]	Distance decay function: linear		Distance decay function: no decay	
		DD_{max} 10% of SP[b]	DD_{max} 50% of SP[b]	DD_{max} 10% of SP[b]	DD_{max} 50% of SP[b]		DD_{max} 10% of SP[b]	DD_{max} 50% of SP[b]	DD_{max} 10% of SP[b]	DD_{max} 50% of SP[b]		DD_{max} 10% of SP[b]	DD_{max} 50% of SP[b]	DD_{max} 10% of SP[b]	DD_{max} 50% of SP[b]
1	4.92	–	–	–	–	13.37	3.76	3.76	3.76	3.76	13.76	4.46	4.46	4.46	4.46
2	6.31	–	–	–	–	27.08	5.50	5.50	5.50	5.50	29.51	4.83	4.83	4.83	4.83
3	11.81	0.68	0.68	0.68	0.68	39.72	4.69	4.69	4.69	4.69	40.92	6.98	7.68	6.98	8.12
4	20.38	–	–	–	–	51.00	4.97	5.06	4.97	5.08	52.76	5.24	8.47	5.38	9.89
5	25.88	1.66	1.66	1.66	1.66	58.56	4.97	5.06	4.97	5.50	61.61	5.96	9.36	6.09	10.85
6	31.26	2.75	2.75	2.75	2.75	63.97	4.11	6.52	4.11	7.64	71.27	3.52	8.06	3.73	10.53
7	37.77	3.64	3.64	3.64	3.64	68.54	3.79	5.36	3.79	6.78	79.04	3.43	9.73	3.65	12.42
8	41.31	3.95	3.95	3.95	3.95	72.32	5.33	9.43	5.54	12.24	85.49	3.12	7.37	3.34	10.12
9	48.76	4.84	4.84	4.84	4.84	77.39	5.25	11.66	5.38	14.79	90.03	2.80	6.10	2.89	7.56
10	50.72	5.25	5.35	5.25	5.36	82.81	7.04	11.42	7.25	13.18	93.99	2.84	4.27	2.84	4.98
11	54.57	5.25	7.03	5.25	7.80	88.75	5.66	8.43	5.66	9.50	95.78	2.03	3.24	2.03	3.76
12	56.26	5.42	7.33	5.44	8.15	94.26	2.54	3.68	2.54	4.50	97.19	1.47	2.32	1.47	2.61
13	57.47	5.25	7.03	5.25	7.80	95.75	2.02	2.76	2.02	3.27	98.32	0.98	1.37	0.98	1.53
14	59.87	5.25	7.03	5.25	7.80	96.73	1.69	2.27	1.70	2.72	99.30	0.55	0.55	0.55	0.66
15	62.64	5.25	7.03	5.25	7.80	97.24	1.48	2.03	1.50	2.47	99.85	–	–	–	0.15
16	64.33	5.42	7.33	5.44	8.15	97.36	2.32	2.40	2.35	2.45	99.93	0.07	0.07	0.07	0.07
17	65.87	5.42	7.33	5.44	8.15	98.05	1.69	1.75	1.72	1.81	100.00	–	–	–	–
18	66.50	5.47	7.49	5.49	8.34	98.18	1.65	1.73	1.68	1.78	100.00	–	–	–	–
19	67.94	5.57	7.00	5.59	7.52	98.21	1.69	1.75	1.72	1.79	100.00	–	–	–	–
20	68.58	5.62	7.16	5.64	7.70	98.33	1.56	1.63	1.59	1.67	100.00	–	–	–	–
21	68.58	5.62	7.16	5.64	7.70	98.33	1.56	1.63	1.59	1.67	100.00	–	–	–	–
22	69.04	5.62	7.16	5.64	7.70	98.33	1.56	1.63	1.59	1.67	100.00	–	–	–	–
23	69.14	5.62	7.16	5.64	7.70	98.33	1.56	1.63	1.59	1.67	100.00	–	–	–	–
24	69.14	5.62	7.16	5.64	7.70	98.33	1.56	1.63	1.59	1.67	100.00	–	–	–	–
25	69.14	5.62	7.16	5.64	7.70	98.33	1.56	1.63	1.59	1.67	100.00	–	–	–	–

"–" Means no improvement compared with the solution of the FRLM with no deviations.

a Percent.

b Gained percent point relative to the FRLM with no deviations.

Table 4 – Optimal coverage gain by indifference to deviation distance.						
p	Range 4		Range 8		Range 12	
	One SP[a]	No decay, 50% SP[b]	One SP[a]	No decay, 50% SP[b]	One SP[a]	No decay, 50% SP[b]
1	4.92	–	13.37	3.76	13.76	4.46
2	6.31	–	27.08	5.50	29.51	4.83
3	11.81	0.68	39.72	4.69	40.92	8.12
4	20.38	–	51.00	5.08	52.76	9.89
5	25.88	1.66	58.56	5.50	61.61	10.85
6	31.26	2.75	63.97	7.64	71.27	10.53
7	37.77	3.64	68.54	6.78	79.04	12.42
8	41.31	3.95	72.32	12.24	85.49	10.12
9	48.76	4.84	77.39	14.79	90.03	7.56
10	50.72	5.36	82.81	13.18	93.99	4.98
11	54.57	7.80	88.75	9.50	95.78	3.76
12	56.26	8.15	94.26	4.50	97.19	2.61
13	57.47	7.80	95.75	3.27	98.32	1.53
14	59.87	7.80	96.73	2.72	99.30	0.66
15	62.64	7.80	97.24	2.47	99.85	0.15
16	64.33	8.15	97.36	2.45	99.93	0.07
17	65.87	8.15	98.05	1.81	100.00	–
18	66.50	8.34	98.18	1.78	100.00	–
19	67.94	7.52	98.21	1.79	100.00	–
20	68.58	7.70	98.33	1.67	100.00	–
21	68.58	7.70	98.33	1.67	100.00	–
22	69.04	7.70	98.33	1.67	100.00	–
23	69.14	7.70	98.33	1.67	100.00	–
24	69.14	7.70	98.33	1.67	100.00	–
25	69.14	7.70	98.33	1.67	100.00	–

"–" Means no improvement compared with the solution of the FRLM with no deviations.
a Percent.
b Gained percent point relative to the FRLM with no deviations.

Table 5 – Optimal coverage gain by multiple shortest paths.						
p	Range 4		Range 8		Range 12	
	One SP[a]	Multi SP[b]	One SP[a]	Multi SP[b]	One SP[a]	Multi SP[b]
1	4.92	–	27.08	3.76	29.51	4.46
2	6.31	–	39.72	5.50	40.92	4.83
3	11.81	0.68	51.00	4.69	52.76	6.98
4	20.38	–	58.56	4.97	61.61	4.71
5	25.88	1.66	63.97	4.97	71.27	4.57
6	31.26	2.75	68.54	4.11	79.04	1.26
7	37.77	3.64	72.32	3.79	85.49	1.17
8	41.31	3.95	77.39	3.07	90.03	1.17
9	48.76	4.84	82.81	4.97	93.99	2.67
10	50.72	5.25	88.75	4.78	95.78	2.84
11	54.57	5.25	94.26	5.66	97.19	2.03
12	56.26	5.25	95.75	2.54	98.32	1.47
13	57.47	5.25	96.73	2.02	99.30	0.98
14	59.87	5.25	97.24	1.63	99.85	0.55
15	62.64	5.25	97.36	1.24	99.93	–
16	64.33	5.25	98.05	1.81	100.00	0.07
17	65.87	5.25	98.18	1.18	100.00	–
18	66.50	5.30	98.21	1.14	100.00	–
19	67.94	5.40	98.33	1.18	100.00	–
20	68.58	5.40	98.33	1.05	100.00	–
21	68.58	5.40	98.33	1.05	100.00	–
22	69.04	5.40	98.33	1.05	100.00	–
23	69.14	5.40	98.33	1.05	100.00	–
24	69.14	5.40	98.33	1.05	100.00	–
25	69.14	5.40	98.33	1.05	100.00	–

"–" Means no improvement compared with the solution of the FRLM with no deviations.
a Percent.
b Gained percent point relative to the FRLM with no deviations.

5.4. Effects of multiple shortest paths

By specifying $DD_{\max} = 0$ (zero), the DFRLM becomes the FRLM with multiple shortest paths (FRLM-MSP). Note that in Table 2 the column of no deviation with a single shortest path indicates the result of the FRLM (FRLM-SSP), while the multiple shortest paths column is for DFRLM with zero deviation distance. One obvious difference of FRLM-MSP solutions is that the computation times are longer than for the FRLM-SSP, which is mainly because of the increased problem size. Specifically, even if no deviation is allowed, the use of a KSP algorithm in generating all deviation paths results in multiple alternative shortest paths for each O–D pair that have the same distance. As a result, the number of paths $|R|$ has increased, which in turn may increase the number of facility combinations $|H|$.

These changes often result in FRLM-MSP reporting a higher objective value even with the same solution set of stations as from the FRLM-SSP. A set of facilities may not provide a feasible combination to refuel one shortest path but this does not necessarily mean that the same set cannot refuel another possible shortest path of the O–D pair. Of course, the solution from FRLM-MSP could be different from the solution of FRLM-SSP. In all cases, FRLM-MSP estimated objective values better than FRLM-SSP (Fig. 10). The gap ranges from 0.07% to 6.98%. An example is the $p = 3$ solution for a range of four (Table 5). FRLM-SSP found 14, 18, and 20 as the solution whereas FRLM-MSP located facilities at 18, 19, and 20 with the higher objective value. The O–D pair of 18–19 can be traversed by nodes 18–17–19 (A) or 18–20–19 (B). Depending on the choice of shortest path generation algorithm and implementation details, FRLM-SSP fixes one from the two paths and generates feasible facility combinations H from it. If the algorithm chose (A) as *the* path, which is the case here, the objective value from the solution of 18, 19, and 20 would be underestimated by not counting in the flows on the path (B) that is exactly as good as (A). That is the reason why FRLM-SSP had to move on to choose 14 whereas FRLM-MSP was able to consider both paths.

6. Conclusions and future research

This paper relaxed the FRLM to consider the willingness of consumers to deviate from their shortest paths to visit a service facility. The number of drivers to visit a facility off of their pre-planned paths was assumed to be decreasing as the required deviation increases. The new model (DFRLM) provides a more realistic representation of refueling behavior especially when the AFV infrastructure is in its infancy. A mixed-integer linear programming formulation was presented and the procedures to generate input data for the model were given. Drivers' sensitivity to deviation can be modeled by one of the procedures. The problem instances were solved to the optimality and the results were discussed.

In general, the results of DFRLM were consistent with the nodes-only version of FRLM. The consideration of deviations in DFRLM resulted in an increase of the objective function even with the same set of facilities as FRLM and this implies higher utilization of facilities and more accurate projection of covered demands. The increase of the objective was consistently observed as the vehicle range or deviation increases given that the objective gain is a composite product of deviation and vehicle range. Depending on the specification of the deviation function, the spatial pattern of optimal stations and the required number of stations to meet a certain level of coverage can change. The MILP formulation of the DFRLM and the procedures for input data generation could be used to solve FRLM problems and more importantly it could enhance the results by eliminating the possibility of coverage underestimation.

This paper contributes to the literature of flow-based location models and models for locating refueling stations by considering deviations for refueling purposes. The structural characteristic of the DFRLM and the procedures for input data generation enhance the results of FRLM by eliminating coverage underestimation. This is achieved by the DFRLM's capability of considering multiple paths between an O–D pair. From a practical point of view, this research provides an important implication for the development of alternative-fuel refueling infrastructure. The results suggested that the choice of deviation decay function and maximum allowed deviation has a measurable effect on solution quality and optimal facility locations. Therefore, careful modeling of deviation behavior in practice is suggested. For example, the infrastructure developers and government agency will need to answer how sensitive potential (and actual) AFV drivers are to the required deviations. Such assessment may be needed in every important phase of the infrastructure development.

This enhancement in reliability and realism required a tradeoff in tractability mainly by needing more time to solve more complex problems. Therefore, improvements in deviation path generation algorithm or development of more flexible feasibility evaluation algorithm will be interesting future research topics. More runs with different deviation behavior models or calibration of parameters with empirical data will provide realism to the DFRLM. Empirical data that can tell us more about AFV drivers are needed. The data will be useful in gathering such important information about AFV drivers as deviation behavior, AFV purchase likelihood, socio-economic characteristics, refueling stops relative to drivers' home and work locations, typical refueling time of a day, or major usage of AFV. This information can be used to calibrate deviation decay model parameters and to realistically estimate geographic variation in AFV demands. The development of a model that can simultaneously consider deviation and capacity is a promising future direction for research. In addition, given that the new model is oriented to be applied to real-world problems, the development of efficient solution approaches for the integrated model is the logical next step. If candidates are not restricted at nodes and they are numerous "enough" relative to vehicle range, the quality of solutions from DFRLM may also increase as in Kuby and Lim [50].

Acknowledgements

We gratefully acknowledge FICO for use of the Xpress Optimization Suite under their Academic Partnership Program.

REFERENCES

[1] Greene DL. Survey evidence on the importance of fuel availability to the choice of alternative fuels and vehicles. Energy Stud Rev 1996;8(3):215–31.
[2] Ogden JM. Developing an infrastructure for hydrogen vehicles: a southern California case study. Int J Hydrogen Energy 1999;24(8):709–30.
[3] Melendez M, Milbrandt A. Hydrogen infrastructure transition analysis: milestone report. Golden, CO: National Renewable Energy Laboratory; 2006. Technical Report No. NREL/TP-540-38351.
[4] National Research Council. The hydrogen economy: opportunities, costs, barriers, and R&D needs. Washington, DC: The National Academies Press; 2004.
[5] Melendez M. Transitioning to a hydrogen future: learning from the alternative fuels experience. Golden, CO: National Renewable Energy Laboratory; 2006. Technical Report No. NREL/TP-540-39423.
[6] Kuby MJ, Lim S. The flow-refueling location problem for alternative-fuel vehicles. Socio Econ Plann Sci 2005;39(2): 125–45.
[7] Berman O, Bertsimas D, Larson RC. Locating discretionary service facilities, ii: maximizing market size, minimizing inconvenience. Oper Res 1995;43(4):623–32.
[8] Hodgson MJ. A flow-capturing location-allocation model. Geogr Anal 1990;22:270–9.
[9] Mirchandani P, Rebello P, Agnetis A. The inspection station location problem in hazardous material transportation: some heuristics and bounds. INFOR 1995;33(2):100–13.
[10] Zeng W, Castillo I, Hodgson M. A generalized model for locating facilities on a network with flow-based demand. Netw Spat Econ 2010;10(4):579–611.
[11] Murray AT, Tong D, Kim K. Enhancing classic coverage location models. Int Reg Sci Rev 2010;33:115–33.
[12] Tester JW. Sustainable energy: choosing among options. Cambridge, Mass: MIT Press; 2005.
[13] Greene DL. TAFV alternative fuels and vehicles choice model documentation. Oak Ridge, TN: Oak Ridge National Laboratory; 2001. Report no.: ORNL/TM-2001/134.
[14] Greene DL, Bowman D. Integrated analysis of market transformation scenarios with hytrans. Oak Ridge, TN: Oak Ridge National Laboratory; 2007. Report no.: ORNL/TM-2007/094.
[15] Melendez M, Milbrandt A. Geographically based hydrogen consumer demand and infrastructure analysis. Golden, CO: National Renewable Energy Laboratory; 2006. Report no.: NREL/TP-540-40373.
[16] Welch C. Lessons learned from alternative transportation fuels: modeling transition dynamics. Golden, CO: National Renewable Energy Laboratory; 2006. Report no.: NREL/TP-540-39446.
[17] Greene DL, Leiby PN, James B, Perez J, Melendez M, Milbrandt A, et al. Hydrogen scenario analysis summary report: analysis of the transition to hydrogen fuel cell vehicles and the potential hydrogen energy infrastructure requirements. Oak Ridge, TN: Oak Ridge National Laboratory; 2008. Report No.: ORNL/TM-2008/030.
[18] Keles D, Wietschel M, Möst D, Rentz O. Market penetration of fuel cell vehicles – analysis based on agent behaviour. Int J Hydrogen Energy 2008;33(16):4444–55.
[19] Ogden JM. Hydrogen delivery model for h2a analysis: a spreadsheet model for hydrogen delivery scenarios. Davis: Institute of Transportation Studies, University of California; 2004. Report No.: UCD-ITS-RR-04-33.
[20] Yang C, Ogden J. Determining the lowest-cost hydrogen delivery mode. Int J Hydrogen Energy 2007;32(2):268–86.
[21] Welch CJ. System dynamics: Hydive(tm) (hydrogen dynamic infrastructure and vehicle evolution) model (presentation); 2007.
[22] Welch CJ. Quantifying consumer sensitivity to hydrogen refueling station coverage. In: Department of Energy, editor. 2007 Annual progress report for the DOE hydrogen program. Department of Energy; 2007. p. 1126–9.
[23] Melaina MW. Initiating hydrogen infrastructures: preliminary analysis of a sufficient number of initial hydrogen stations in the US. Int J Hydrogen Energy 2003; 28(7):743–55.
[24] Melaina MW. Estimating relative station sizes in early hydrogen station networks. In: Proceedings of the National Hydrogen Association, Washington, DC; 2005.
[25] Melaina MW, Bremson J. Regularities in early hydrogen station size distributions. In: Proceedings of the 26th United States Association for Energy Economics. Ann Arbor, Michigan: International Association for Energy Economics North American Conference; 2006.
[26] Melaina M, Bremson J. Refueling availability for alternative fuel vehicle markets: sufficient urban station coverage. Energy Policy 2008;36(8):3233–41.
[27] Kitamura R, Sperling D. Refueling behavior of automobile drivers. Transp Res Part A Policy Practice 1987;21(3):235–45.
[28] Ni J, Johnson N, Ogden JM, Yang C, Johnson J. Estimating hydrogen demand distribution using geographic information systems (GIS), Presentation series. Davis, CA: Institution of Transportation Studies; 2005.
[29] Melendez M, Milbrandt A. Analysis of the hydrogen infrastructure needed to enable commercial introduction of hydrogen-fueled vehicles. Golden, CO: National Renewable Energy Laboratory; 2005. Report No.: NREL/CP-540-37903.
[30] Hakimi SL. Optimum locations of switching centers and the absolute centers and medians of a graph. Oper Res 1964; 12(3):450–9.
[31] Toregas C, Swain R, ReVelle CS, Bergman L. The location of emergency service facilities. Oper Res 1971;19(5):1363–73.
[32] Church RL, ReVelle CS. The maximum covering location problem. Pap Regional Sci 1974;32:101–18.
[33] Hakimi SL. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. Oper Res 1965;13(3):462–75.
[34] ReVelle CS, Swain R. Central facilities location. Geogr Anal 1970;2:30–42.
[35] Balinski ML. Integer programming; methods, uses, computation. Manage Sci 1965;12:253–313.
[36] Kuby MJ. Programming models for facility dispersion: the p-dispersion and maxisum dispersion problems. Geogr Anal 1987;19:315–29.
[37] Bapna R, Thakur LS, Nair SK. Infrastructure development for conversion to environmentally friendly fuel. Eur J Oper Res 2002;142(3):480–96.
[38] Bersani C, Minciardi R, Sacile R, Trasforini E. Network planning of fuelling service stations in a near-term competitive scenario of the hydrogen economy. Socio Econ Plann Sci 2009;43(1):55–71.
[39] Huff DL. Defining and estimating a trade area. J Marketing 1964;28:34–8.
[40] Goodchild M, Noronha V. Location–allocation and impulsive shopping: the case of gasoline retailing. In: Ghosh A, Rushton G, editors. Spatial analysis and location–allocation models. New York: Van Nostrand Reinhold; 1987. p. 121–36.
[41] Chan TY, Padmanabhan V, Seetharaman PB. An econometric model of location and pricing in the gasoline market. J Market Res (JMR) 2007;44(4):622–35.
[42] Lin Z, Ogden J, Fan Y, Chen C-W. The fuel-travel-back approach to hydrogen station siting. Int J Hydrogen Energy 2008;33:3096–101.

[43] Nicholas MA, Handy SL, Sperling D. Using geographic information systems to evaluate siting and networks of hydrogen stations. Transp Res Record J Transp Res Board 2004;1880:126–34.
[44] Hodgson MJ, Rosing KE, Zhang J. Locating vehicle inspection stations to protect a transportation network. Geogr Anal 1996;28(4):299–314.
[45] Hodgson MJ, Berman O. A billboard location model. Geogr Environ Modell 1997;1:25–45.
[46] Berman O, Larson RC, Fouska N. Optimal location of discretionary service facilities. Transp Sci 1992;26(3):201–11.
[47] Wang Y-W, Lin C-C. Locating road-vehicle refueling stations. Transp Res Part E Logistics Transp Rev 2009;45(5): 821–9.
[48] Zeng W, Hodgson MJ, Castillo I. The pickup problem: consumers' locational preferences in flow interception. Geogr Anal 2009;41(1):107–26.
[49] Berman O, Hodgson JM, Krass D. Flow-interception problems. In: Drezner Z, editor. Facility location: a survey of applications and methods. New York: Springer; 1995. p. 389–426.
[50] Kuby MJ, Lim S. Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. Netw Spat Econ 2007;7(2):129–52.
[51] Upchurch C, Kuby MJ, Lim S. A model for location of capacitated alternative-fuel stations. Geogr Anal 2009;41: 85–106.
[52] Lim S, Kuby MJ. Heuristic algorithms for siting alternative-fuel stations using the flow-refueling location model. Eur J Oper Res 2010;204(1):51–61.
[53] Capar I, Kuby MJ, Rao B. A new formulation of the flow-refueling location model for alternative-fuel stations. IIE Transactions, in press.
[54] Kuby MJ, Lines L, Schultz R, Xie Z, Kim J-G, Lim S. Optimization of hydrogen stations in Florida using the flow-refueling location model. Int J Hydrogen Energy 2009;34(15): 6045–64.
[55] Hodgson MJ. The location of public facilities intermediate to the journey to work. Eur J Oper Res 1981;6(2):199–204.
[56] Hoffman W, Pavley R. A method for the solution of the nth best path problem. J ACM 1959;6(4):506–14.
[57] Yen JY. Finding the *k* shortest loopless paths in a network. Management Sci 1971;17(11):712–6.
[58] Dreyfus SE. An appraisal of some shortest-path algorithms. Oper Res 1969;17(3):395–412.
[59] Shier DR. On algorithms for finding the k shortest paths in a network. Networks 1979;9(3):195–214.
[60] Eppstein D. Finding the *k* shortest paths. SIAM J Comput 1994;28(2):652–73.
[61] Berman O, Simchi-Levi D. A heuristic algorithm for the traveling salesman location problem on networks. Oper Res 1988;36:478–84.
[62] ReVelle C. Facility siting and integer-friendly programming. Eur J Oper Res 1993;65(2):147–58.