Geospatial, Temporal and Economic Analysis of Alternative Fuel Infrastructure: The Case of Freight and U.S. Natural Gas Markets

Yueyue Fan,[a] *Allen Lee,*[b] *Nathan Parker,*[c] *Daniel Scheitrum,*[d]* *Amy Myers Jaffe,*[e]
Rosa Dominguez-Faus,[e] *and Kenneth Medlock III*[f]

ABSTRACT

The transition to low-carbon fuel in the United States has spatial, temporal and economic aspects. Much of the economic literature on this topic has focused on aspects of the cost effectiveness of competing fuels. We expand this literature by simultaneously considering spatial, temporal and economic aspects in an optimization framework that integrates geographic information system (GIS) tools, network analysis, technology choice pathways and a vehicle demand choice model. We focus on natural gas fuel as a low-carbon alternative to oil-based diesel fuel in the heavy-duty sector primarily because of the recent cost benefits relative to diesel fuel and the high vehicle turnover rate in heavy-duty trucks. We find that the level of profitability of natural gas fueling infrastructure depends more on volume of traffic flows rather than proximity to natural gas supply.

Keywords: Natural gas, Spatial optimization, Alternative fuels, Fuel infrastructure

https://doi.org/10.5547/01956574.38.6.yfan

1. INTRODUCTION

The recent emergence of natural gas as an abundant, inexpensive fuel in the United States could prompt a momentous shift in the level of natural gas utilized in the transportation sector. The cost advantage of natural gas vis-à-vis diesel fuel is particularly appealing for vehicles with a high intensity of travel and thus fuel use, such as long distance trucking. Natural gas is already a popular fuel for short-haul, urban municipal, and fleet vehicles such as garbage trucks,[1] transit buses, and taxis. The so-called "shale revolution" has unleashed a giant surge in U.S. natural gas production that is making natural gas a competitively priced fuel in many different applications (Medlock, 2012).

1. Only one automobile manufacturer, Honda Motor Co., had been offering a natural gas passenger vehicle for sale in the United States, but the car has since been taken off the market.

[a] University of California-Davis, Civil and Environmental Engineering.
[b] Data Scientist, DogVacay.com.
[c] Arizona State University, School of Geographical Sciences & Urban Planning.
[d] University of California-Davis, Department of Agricultural and Resource Economics.
[e] University of California-Davis, Institute for Transportation Studies.
[f] James A. Baker III Institute for Public Policy, Rice University.
* Corresponding author. E-mail: dpscheitrum@ucdavis.edu.

Long distance freight could be the next logical place for natural gas to expand penetration. The long distance freight market would be a beneficial place for policy makers to target alternative fuels use because the high turnover rate for heavy-duty trucks (roughly 200,000 to 240,000 new heavy-duty vehicles come on the road each year) means that steady demand for new trucks could be a facilitating factor. Port cities that are serviced by these fleets face air quality problems that could be ameliorated by cleaner fuel, adding to the advantages. Between 2014 and 2025 roughly 2.7 million new trucks will be purchased—or 76% of the total fleet in 2025, creating a ready market for natural gas vehicles. At present only 14% of fleets operate any vehicles on alternative fuels (National Petroleum Council, 2012; Straight, 2014).

To gain significant market share, natural gas fueling infrastructure would have to compete against a mature, robust distribution network for diesel fuel. There are 59,739 diesel fueling stations and 121,446 gasoline stations across the United States and 2,542 truck stops where fuel is readily and conveniently available. By contrast, there are just 800 CNG fueling sites, and just under half are available to the public (Alternative Fuels Data Center, 2017a). There are currently 59 open-to-the-public LNG fueling stations and 42 private LNG fueling stations along routes from Los Angeles to Las Vegas, around Houston and around Chicago. The stations currently serve a fleet of 3,600 LNG trucks (Alternative Fuels Data Center, 2017b). Zeuss Intelligence reports that there are 34 LNG supply plants with trailer loadout capable of producing about 3 million gallons of LNG a day.[2]

We choose natural gas as a proxy for alternative fuels[3] to highlight factors other than relative fuel cost because natural gas requires new fueling infrastructure and vehicle stocks like many alternative fuels but is less expensive than diesel fuel, and engine and fuel dispensing technologies are already proven. This simplifies the number of factors that are salient and allows researchers to consider more directly many other variables influencing the adoption rate of alternative fuels other than fuel cost or technological risk.

In this paper, we investigate the possibility that natural gas could be utilized to provide fuel cost savings and geographic supply diversity for the heavy-duty trucking sector and whether it can enable a transition to lower carbon transport fuels. To answer the question about the optimal locations for natural gas fueling, we use a modeling framework that utilizes spatial mapping of existing major interstate highways, trucking routes, key fueling routes for fleets and heavy-duty trucks, and fueling infrastructure for CNG for fleet operation and LNG for long-haul trucks to make infrastructure planning decisions. Spatial network theory and network analysis are used to calculate the most profitable trucking corridors to establish LNG infrastructure (Lee et al., 2015). We consider whether the wide availability of surplus localized natural gas supplies is a major feature that promotes adoption of a natural gas-fueled trucking network. A number of states are offering incentives to support the expansion of natural gas as a transportation fuel, including Pennsylvania, home to the rich Marcellus shale gas basin, and Oklahoma, which also has prolific resources. Our investigation posits the question of whether the rate of vehicle flow over a route or its proximity to supplies creates the most powerful driver to profitability and thereby network expansion.

2. Provided to authors by Zeuss Research consultants.

3. Alternative fuels have been widely covered in the academic literature. Our research builds on this literature but allows for a more focused approach that simplifies the vast array of features that face those other fuels to highlight structural issues related to the incumbent industry. On alternative fields we recommend Bandivadekar et al. (2008), Greene et al. (2008), and Ogden and Nicholas (2011). For an example of work on this topic see Apanel and Johnson (2004), Dominguez-Faus et al. (2013), Knittel (2012), Morrison et al. (2014), and Ogden et al. (1999).

We expand existing literature on refueling network research for planning and designing future hydrogen supply chains to improve knowledge related to transportation applications for natural gas and thereby aim to elucidate additional knowledge about barriers to alternative fuels adoption rates (Dagdougui, 2012). We utilize an optimization solution that contributes to existing modeling science by including network analysis to a modeling solution that simultaneously considers temporal, spatial, technological and economic aspects (Dagdougui, 2012). Our paper extends work by Parker et al. (2010) and Kuby et al. (2009) and allows for the direct comparison of economic and cost considerations against the spatial requirements of the entire supply chain and refueling network. Our optimization solution further contributes to existing modeling science by applying network analysis to reduce the number of potential candidate locations for new infrastructure from infinitely many to a finite number of reasonable choices given. We combine and augment these approaches in our study to take into account spatial data of existing infrastructure as well as the economic and technological requirements of the supply chain and to evaluate inter-temporally a buildout solution over a multi-period time horizon. Our methodology allows for the comparison of supply chain cost considerations against spatial criteria to reveal the most important determinants of alternative fuel refueling development. This comparison was not possible with the generalized mathematical optimization method or spatial analysis method used alone. Utilizing network analysis inside this general modeling framework, our aim is to determine the most profitable transportation networks and locations for natural gas flows into transportation markets nationwide.

Our model uses the spatial infrastructure data and compares costs for transportation of natural gas by source, distribution method, and other market development variables through mathematical optimization. In other words, we study where the most cost-effective and profitable locations to build natural gas fueling infrastructure would need to be located in order to minimize the costs of an overall national system of natural gas fueling along major inter-state highways. A previous study by Rood Werpy concludes that high costs, limited refueling infrastructure, and uncertain environmental performance constitute barriers to widespread adoption of natural gas as a transportation fuel in the United States (Werpy et al., 2010). However, in another substantial contribution to the literature, Krupnick finds that the move from a long-haul route structure to a "hub and spoke" structure could facilitate the development of natural gas refueling infrastructure in the highway system (Krupnick, 2011). Our study is the first to utilize supply chain optimization techniques with network spatial analysis and link to a simplified natural gas demand model to explore and analyze natural gas infrastructure for the on-road heavy-duty and freight transportation sector.

2. METHODOLOGY, MATHEMATICAL FORMULATION OF THE MODEL AND POTENTIAL TECHNOLOGY PATHWAYS

For natural gas to be successful in the heavy/medium-duty fleet market, there must be an integrated network of public access stations and natural gas infrastructure across the country that can support significant penetration of natural gas vehicles for long distance travel. In this study, we will focus specifically on liquefied natural gas (LNG). Successful LNG infrastructure implementation seeks to minimize one or more of the four main cost components of the LNG supply chain: feed-gas cost, liquefaction, transportation, and refueling (TIAX for America's Natural Gas Alliance, 2010). For the modular LNG pathway, feed-gas cost represents 24% of total cost, transportation costs compose 7%, and station capital and operating costs compose 69%. For the conventional pathway, feed-gas cost composes 14% of total cost, liquefaction costs makes up 25%, transportation 22%, and station capital and operating costs make up 39% of total cost. These cost shares are consistent across stations, though the conventional pathway costs are slightly more variable. Planning a profitable and sustainable natural gas fueling infrastructure requires careful selection of station locations and their capacities to allow maximum usage of existing facilities and newly built stations. Strategic and coordinated investments along heavily used truck corridors, such as the establishment of co-located natural gas stations and diesel truck stops, will be required to establish infrastructure networks that make LNG a major player. Therefore, in creating a modeling framework to study the potential economics of a national American natural gas fueling network, we investigate the entire LNG supply chain, instead of treating different facilities separately, to achieve the best system performance while satisfying current demand and evolving to accommodate incremental future demand.

2.1 Literature Review

It has been widely recognized that the performance of an energy supply system depends on the setting of the entire supply chain instead of any individual technology or facility (Parker et al., 2008). By and large, alternative fuel research has focused on hydrogen-based systems, as evidenced by a recent review, and revolves around refueling network research for planning and designing future hydrogen supply chains (Dagdougui, 2012). There are a small number of studies that combine the strengths of a national energy system optimization approach with a spatially explicit infrastructure optimization approach. Stachan et al. described an integrated approach linking spatial GIS modeling of hydrogen infrastructures with an economy-wide energy systems (MARKAL) model's supply and demand (Strachan et al., 2009). Parker et al. (2010) used an annualized profit maximization formulation to study the optimal distribution network for bio-waste to hydrogen. Kuby et al. applied a flow-capturing location model to optimizing the locations of hydrogen stations in Florida using real world traffic and demographic data (Kuby et al., 2009). The prevailing focus of research on product innovation for alternative fuels in heavy-duty trucking has been on natural gas, though recent work by Fulton and Miller is extending this body of research to consider electrification as well (Fulton and Miller, 2015).

The recent introduction of LNG fuels into the heavy-duty vehicle sector has received minimal academic attention despite a growing interest for the fuel as an alternative to diesel. It has been established that the cost of developing LNG infrastructure may be what is hindering the deployment of heavy-duty natural gas vehicles (HDNGVs) (National Renewable Energy Laboratory, 2015). In this paper, we propose methodology involving the LNG supply chain to assist in the exploration of both the barriers and solutions to commercializing LNG for the transportation sector. A review of current LNG supply chain literature shows that applications are rather limited and only available for the maritime industry, involving the transshipment of LNG overseas (Kristoffersen and Fleten, 2010; Grønhaug and Christiansen, 2009). This paper attempts to shed light into a new application for LNG supply chains—LNG as a domestic on-road transportation fuel for HDNGVs. To the best of our knowledge, this will be the first study that utilizes supply chain optimization techniques with network spatial analysis and links it to a simplified LNG demand model to explore and analyze LNG infrastructure for the on-road heavy-duty transportation sector.

At the heart of the supply chain optimization model is the facility location model. The general facility location problem involves a set of spatially distributed customers and a set of facilities to serve customer demands (Hamacher, 2002). Such problems can be classified into two broad categories based on the relationship between demand and facility, those being point-based demand and flow-based demand. The majority of facility location research has been devoted to designing models that optimally service point-based demands. Examples include the p-median

(Hakimi, 2014)(Charles S. Revelle, 1970), location set cover (Constantine Toregas, Ralph Swain, 1970), maximum cover (Church and Revelle, 1972), and the fixed-charge (Balinski, 1965) models. All four models assume that demand is expressed at point locations where the important distance is the direct path from point-demand to facility. The less popular flow-based demand approach (Agnolucci and McDowall, 2013; Balinski, 1965; Kuby and Lim, 2005; Zeng et al., 2010), asserts that demand behaves more like traffic flows on an origin-destination (O-D) route and that the goal is to capture as much traffic volume as possible. An example of such model is the flow capturing location model (FCLM) by Hodgson (Kuby and Lim, 2005). It is easiest to understand these two different approaches if we introduce the notion that flow-based demand is "refueling on the go" as opposed to "home to station or work to station refueling," which is the equivalent concept for point-based demand, typically applied to light duty vehicles. In other words, flow-based models assume that vehicles tend to refuel en route to their destination rather than making special purpose trips from home or work to the gasoline station (Kuby and Lim, 2005). The concept of flow-based demand closely resembles the behavior of heavy-duty freight trucking operations because truckers refuel en route to their destination. Therefore, we use flow-based facility location model in this study.

Another important aspect of supply chain modeling regards consumer demand. As pointed by Agnolucci (Agnolucci and McDowall, 2013), realistic representation of demand is perhaps the most critical and arguably the most difficult to model in supply chain optimization models. Since most supply chain networks are demand-driven (Murthy Konda et al., 2011), it is important to have an accurate representation of both spatial and temporal demand variations. Most existing studies focus on point-based demand. For example, Ni et al. (2005) used demographic variables to identify demand centers on the basis of population density, car ownership, market penetration rate, and fuel use but mostly population density. The U.S. National Renewable Energy Laboratory (NREL) conducted a study to identify favorable hydrogen adoption regions (Lin et al., 2008). Johnson et al. (2008) performed a regional GIS-based assessment of coal-based hydrogen infrastructure deployment for the state of Ohio where they used census-based population densities to derive hydrogen demand at clusters. As for flow-based demand assessment, Lin (Lin et al., 2008a; Lin et al., 2008b) utilized vehicles miles travelled (VMT) to represent consumer demand. Kuby et al. (2009) used a gravity model to estimate intercity flows based on zone population and a friction function derived from Michigan's statewide travel forecasting model. Kelley and Kuby (2013) find that CNG vehicle owners refuel on the way ten times more often than refueling at home and suggest flow-capturing models are superior to point-demand models in the natural gas sector. In this study, we use data from the Commodity Flow Survey (CFS), which provides survey data for vehicle tonnage transported between major freight corridors in the U.S. With CFS, we are able to derive the directional traffic volumes on each major corridor route and as a result, estimate regional demand by trucking route.

The temporal dimension of demand has not been well integrated to transportation energy supply chain models. In a review of general facility location and supply chain management, Melo et al. (2009) found that 82% of surveyed papers presented a single-period problem to begin with, leaving temporal variation out of the question. Most national scale energy system optimization models that incorporate temporal demand variation do not have a spatial structure (Agnolucci and McDowall, 2013). There are a small number of studies that combine the strengths of a national energy system optimization approach with a spatially explicit infrastructure optimization approach, typically via coupling MARKAL energy systems model with a single-stage spatial energy infrastructure model and iteratively passing results between the two models (Strachan et al., 2009) (Rosenberg et al., 2010).

2.2 Modeling Framework

To analyze the potential for an expansion of natural gas into the heavy-duty trucking sector and its widespread use across the country's major trucking routes, we create a modeling framework that utilizes spatial mapping of existing major interstate highways, trucking routes, key fueling routes for fleets and heavy-duty trucks, and fueling infrastructure for LNG for long-haul trucks to make infrastructure planning decisions. Spatial network theory and network analysis is utilized to generate all of the spatial information that is needed to calculate the most profitable trucking corridors to establish LNG infrastructure (Lee et al., 2015).

Our spatial optimization model is designed to determine the most profitable transportation networks and locations for natural gas flows into transportation markets nationally using the spatial infrastructure data and comparing costs for transportation of natural gas by source, distribution method, and other market development variables through mathematical optimization.

Our study considers where and when infrastructure should be deployed over a 20-year time horizon in order to satisfy demand along major trucking routes or corridors. A mixed-integer linear programming model is developed to optimize the process of building and operating LNG liquefaction and distribution facilities during the transition to using more LNG in the heavy-duty vehicle sector. We solve our model at roughly five-year increments from 2012 to 2030.[4] We take a rolling window approach, i.e., at each decision stage, we update the station specific LNG demand and then solve for the optimal incremental station construction and existing station upgrades taking the current state of the LNG infrastructure as given during the previous time period.

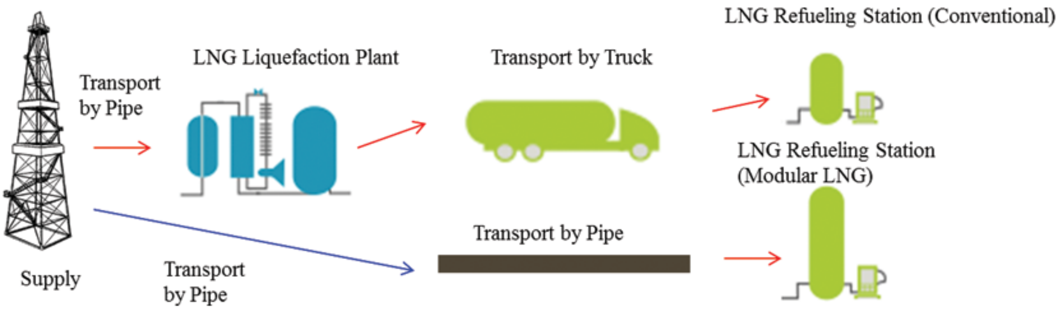
In the optimization model, when selecting refueling station sites from a candidate pool, we include the constraint that stations along a route are only constructed if the route taken as a whole is profitable and therefore if the route can be commercially feasible. The profitability constraint does not require that each station is profitable, but rather if there are unprofitable stations along a route, the more profitable stations must earn enough to compensate for the stations operating at a loss. We are considering route-level profitability on the basis that a single developer could construct all stations along a route. This is not entirely unrealistic since in our optimization model, a route can consist of as few as three stations. A single developer could construct two profitable stations at the end points plus one station operating at a loss in the middle, but the unprofitable station is necessary to allow truck to travel from end to end and therefore complete the route. It is worth pointing out that in the current market development in the United States of natural gas fueling for heavy trucking, there are many routes that have been developed by a single developer for a region (for example, ENN Group Co. Ltd. in Utah) or even national routes in the case of Clean Energy's America's Natural Gas Highway. Since we are not assuming market-power, solving the objective function from a route-level basis can be seen as the socially optimal configuration. We acknowledge that while the assumption of single-owner development does limit the models widespread application, the results of this model present a bound on infrastructure deployment.[5] Future use of this model aimed outside the long-haul heavy duty truck market will require significant modification to allow for the strategic interaction of multiple possible entrants.[6] The issue of non-

4. We solve our model for the years 2012, 2015, 2020, 2025, and 2030.

5. For example, for a max-profit infrastructure development problem, a single-owner model (also called "system-optimal" model) will result in a total profit that is higher than what could be achieved in a non-cooperative multi-owner model (also called "user-optimal"), thus setting an upper bound of the system-wide performance.

6. For instance, modeling of urban delivery and transport markets will require the integration of strategic interaction of multiple entrants.

Figure 1: Natural Gas Refueling Pathways



cooperative refueling/charging infrastructure investment is addressed by Guo et al (2016) and can provide a basis for extension of this model.

In addition, one unique feature that we must address is to explicitly consider vehicle range. LNG fuel contains roughly 60% of the energy density of diesel fuel. Additional storage tanks can be added but this is unlikely since that would take cargo space and storage technology is expensive. This makes LNG trucks range limited when compared with traditional diesel trucks. The feasibility constraint requires that stations must be no further apart than the maximum range of LNG trucks. It is this feasibility constraint that sometimes leads to the construction of unprofitable stations in order to ensure that a route can be traversed by LNG vehicles. However, the profitability constraint requires that the losses incurred by unprofitable stations are at least offset by profits at more successful stations. This choice is designed to mirror the likely investment decision-making profile of potential station investors. Investors while cognizant that customers will not switch to LNG unless a fully-built out fueling infrastructure exists along their entire operational route, will still require to be assured of an overall return to capital for infrastructure implementation. We assume a rate of return to capital of 12% in line with common commercial practices.

We consider the economics of two alternative LNG delivery routes shown in Figure 1. The first delivery route is the conventional LNG pathway. In the conventional pathway, natural gas is delivered from the supply site to a liquefaction plant via pipeline. After it is liquefied, it is delivered by truck to a refueling station and put into a storage tank.[7] LNG is then dispensed out of the storage tank at the refueling station. The second delivery route is the modular LNG pathway. In this pathway, natural gas is delivered from the supply site directly to the refueling station via pipeline. At the refueling station, natural gas is then converted to LNG onsite in a modular liquefaction plant. The LNG is then dispensed to the customer.

At each candidate refueling station site, we consider the cost of dispensing LNG on a per gallon basis taking into account station capital cost, operating and management cost, feed-stock cost, fuel transport cost (feedstock pipeline transport cost plus trucking cost if conventional pathway), pipeline construction cost, and lastly, electricity cost. At each candidate site, we select the technology and capacity that minimizes the delivered LNG per gallon price. This comprehensive assessment tool is aimed to simulate the potential volumetric capacity for the natural gas transpor-

7. We assume that trucking costs are $10/mile per truckload with a typical truckload equal to 12,420 LNG gallons. Therefore, we use $0.00085 per mile per gallon as our estimate cost of delivering LNG by truck.

tation market in the United States, as well as to choose the optimal location of new and existing fueling facilities subject to this simulated volumetric capacity.

Production of LNG has strong economies of scale in both capital and operating costs. At present, companies are studying the trade-offs between configurations that can reduce the capital expense at the cost of higher operating costs (modular LNG) versus configurations with higher initial capital expenses but lower operating costs (conventional LNG and trucking). In addition, the price and availability of natural gas across the United States is not spatially uniform and there may be some cost advantages to developing LNG fueling infrastructure near existing LNG receiving/export plants in key coastal markets. The spatial dimension of the system causes tradeoffs between the capital cost of plants and the transportation cost of LNG, which need to be considered in the planning process.

The total costs for each technology pathway are compared in the optimization solution. In the case of conventional LNG, it is assumed that LNG fuel produced at the plant must then be shipped by truck.[8] Commercial economics restricts that network to a 350 square mile radius from the plant. Modular LNG requires access to a natural gas pipeline supply. The supply chains for each of the possible LNG technologies are presented in Figure 1.

Our basic supply chain model of the optimal design of LNG infrastructure captures the decisions of which truck stops to retrofit with LNG availability and whether to create LNG on-site with modular units or through addition of capacity at a centralized production facility with truck delivery. The solution considers the optimal location and capacity of centralized production facilities and refueling stations.

The objective function of the optimization model is to maximize the total system profits from the point of view as if there was a single owner of all stations operating without market power, subject to a series of constraints. Our optimization maximizes profits of the total system on a route-by-route basis accounting for all supply chain costs, subject to latent trucking LNG demand and subject to profitability and feasibility constraints. The revenue stream is estimated based on the price of diesel at the time of writing. The natural gas costs include the following: fixed station cost (includes cost of building additional pipeline), variable station cost (based on throughput capacity), pipe transport cost (service cost per unit flow), fixed liquefaction plant cost, variable plant cost (based on throughput capacity), liquefaction plant transport cost (service cost to get gas from supply source to liquefaction plant), and trucking LNG transport cost (service cost to get LNG from liquefaction plant to LNG truck stop).

The model chooses the optimal locations, technologies, and capacities for LNG stations from a set of candidate locations based on existing diesel truck stop locations.[9] A binary decision making process will decide whether or not to construct a natural gas fueling option at any given diesel truck stop location. The model selects the type of LNG station based on the optimal options for liquefaction plants from a set of candidate locations and related fueling station locations within its commercially profitable sphere of operation.[10] The model implicitly decides the extent to which natural gas is piped and LNG trucked at any point in the supply chain in order to maximize total

8. LNG must be kept extremely cold, -162° C. Delivery by truck allows for the LNG to be kept cold by storing it in a specialized, highly-insulated container. Delivery of LNG by pipeline is not feasible due to the limitations of cooling an insulating an entire length of pipeline.

9. In real world practice, safety and other commercial land-use and zoning constraints will also dictate how close to existing diesel fueling infrastructure natural gas fueling can be built.

10. In the case of conventional LNG trucking routes, we assume a 350-mile radius from the plant location as a commercial constraint based on industry economics and practice.

system profits. Distance between stations considers the maximum practical fuel range of heavy duty vehicles including a small tank reserve for smooth operations of LNG fuel. In addition, the model solves for the optimal flows between supply sources and refueling stations.

Ideally, we would like to include all existing diesel truck stops and diesel supply infrastructure facilities in the candidate location pool. We also include existing petroleum terminal locations as candidate locations for liquefactions plants. However, the size of the problem would be too large to keep a reasonable computational tractability for our model. To form a reduced-size candidate location pool, we first used a spatial analysis (McHarg, 1995) tool in GIS to select candidate locations based on real industry input to include facility locations within certain proximity of a major interstate or pipeline. Then using classical centrality theory of betweenness centrality metrics (Freeman, 1977) and Urban Network Analysis tool (Sevtsuk and Mekonnen, 2012), we were able to identify critical nodes and links to be included in the candidate pool for infrastructure placement. After the first two steps, there were still many existing diesel facilities not included in the candidate pool. In order to account for those remaining facilities yet keeping a reduced-size candidate pool, we used *k*-means clustering algorithms within GIS for those remaining facilities and included the centroids of the clusters as candidate locations in the pool. Such treatment allows us to keep a balance between spatial details and computational tractability for our model.

2.3 Natural Gas Trucking Demand Model

To determine the amount of new demand for natural gas fuel that can emerge in any given year, we develop a trucking demand model to represent the constraints on demand growth through the natural rate of truck turnover and the economic competitiveness of LNG across the distribution of truck use. The demand for LNG from trucking is tied to the turnover of trucks in the market and the competitiveness of LNG as a fuel including both the cost of the truck and the cost of the fuel. The demand model represents the latent demand that could develop if the fueling infrastructure is created to serve it.

The trucking demand model first models fleet turnover. We identify the age distribution of long-haul trucks in operation as of 2012 and based on historical truck survival rates, calculate the number of trucks that will be scrapped each year into the future. Using as sales/scrap ratio of 1.1 through 2020 and a sales/scrap ratio of 1.2 for 2021 to 2030, the model yields the number of projected new long-haul trucks through 2030. For new truck sales, we model the choice to opt in to natural gas technology as whether the fuel price discount of natural gas relative to diesel is sufficient to payback the incremental cost of a natural gas truck within three years as a function of the annual mileage. Truck with greater annual mileage will enjoy greater fuel cost savings and will be more likely to be able to recoup the additional cost of a natural gas truck within three years. Lastly, as a function of the fuel price discount of natural gas relative to diesel, we report the percent of total annual truck vehicle miles traveled that are travelled by natural gas trucks (the penetration rate) each year. A more comprehensive explanation of the truck demand model is reported in Appendix A.

2.4 Data Sources

We have compiled the necessary data for this project from a variety of sources. The data on annual average truck traffic for the United States was obtained from the Bureau of Transportation Statistics and Freight Analysis Framework. We rely on the Federal Highways Network per the Bureau of Transportation Statistics for data on the location of United States Highways. For the location of natural gas pipelines, we rely on the National Natural Gas Pipeline Network from National Pipeline Mapping System. Data on locations of natural gas trading hubs and natural gas spot prices were obtained from the Baker Institute World Gas Trade Model (Medlock and Hartley, 2006). The price of diesel by state was obtained from the United States Energy Information Administration. The price of electricity used as an input for liquefaction is from the United States Energy Information Administration. Our source for LNG refueling station candidate locations is the actual locations of diesel refueling stations in the United States from a commercial dataset called "Diesel Truck Stops."[11] Our source for LNG liquefaction plant candidate locations is calculated from National Natural Gas Pipeline Network as described in section 2.2 and in Appendix B. Our information on LNG refueling station costs and LNG liquefaction plant costs comes from survey of industry including General Electric, Cummins, and Westport.

2.5 Station Demand Calculations

Station LNG demand is estimated by trucking route demand. Data from the Bureau of Transportation Statistics and Freight Analysis Framework has collected commodity flow data from major trucking and shipment companies in between almost every major U.S. city. Using this commodity flow data, we convert it into the annual average daily truck traffic or in other words the number of trucks traveling between its origin and its destination. For the model to opt whether to build a route, the stations on that route must be able to satisfy the fuel requirements of natural gas trucks travelling along the route. In order to manage computational burden, we used a reduced set of routes classified as having 90% of all truck traffic demand.

The demand model is linked with the optimization model through a lookup table of updating penetration rates based on a given fuel price difference produced by the optimization model. A new separate penetration rate is assigned to each station, and the optimization model is again resolved given this new piece of data.

The metric used to determine potential profitability for a station is simple—it is the price difference between the retail price of diesel and the model's calculated price of LNG in equivalent units or $/mile.[12] For example, if the price difference is positive, it means diesel is more expensive, and vice versa. The larger the price differences the better from the point of view of fuel switching. The larger the price difference, the more favorable that station will appear to truckers and thus more demand it will receive.

Every station location is associated with a quantity of truck vehicle miles associated with that location derived from the national truck traffic data. From this associated vehicle miles traveled, we compute LNG demand at the station location by considering the natural gas penetration rate and average efficiency of the vehicles. That is the percentage of vehicle miles travelled that are attributed to LNG vehicles. In the first stage, we choose the initial penetration rate and apply it to all stations in the nation.[13] In later stages, the penetration rate at each station is updated dynamically based on feedback from the trucking demand model which determines the proportion of new vehicles that would choose LNG technology based on a 3-year payback rule given the LNG-Diesel price differential at a given station. If a station has a zero or unfavorable "price differential," *i.e.*

11. The data set was donated to UC Davis.

12. Input cost for feed-gas is based on available cost of gas at the closest pipeline hub plus transportation costs.

13. We consider many choices of initial penetration rate to evaluate the sensitivity of the model and ideal conditions that yield a national roll-out

profitability, the penetration rate remains the same until the next stage of possible expansion in demand in five years. It is this dynamic demand feedback that allows the network to evolve and expand coverage over our time-horizon.

We create a flow-based algorithm similar to the one created by Kuby and Lim (2005) that will generate all combinations of refueling stations able to refuel a specific origin-destination truck route that comply with the feasibility constraint. Reducing the set of station combinations to only those that can allow a range-limited vehicle feasibly traverse a route aids in greatly reducing the number of calculations required. The details of preprocessing the set of station combinations to retain only the feasible subset are provided in 2.2 and Appendix B.

2.6 LNG Supply Chain Optimization Model

At each decision stage, we developed an LNG supply chain optimization model that incorporates explicit spatial and temporal variations in supply and demand, competition between different technology pathways, and economies of scale comparisons in finding the best configuration for the LNG supply chain. The model locates, sizes, and allocates LNG to refueling stations with the objective of maximizing the profitability of the industry as a whole. The profit considered is the sum of the profits for each individual station owner over the entire study region. Costs considered are those associated with supply procurement, transportation, and fixed and variable technology costs. Fuel production and selling price determine the industry revenue.

The supply chain optimization model can be categorized as a deterministic, multi-commodity, capacitated facility location problem that is formulated as a mixed integer linear program, which solves for which truck route, (W_r) to deploy over the course of the study horizon. The model solves for which particular feasible station combination, (V_h) will be built. Moreover, this station combination will help determine the LNG liquefaction plants l and at what size s should be built, (Y_{ls}); and which LNG refueling stations j of technology t should be build, (X_{jt}); and if built, how much natural gas in LNG gallons should be delivered between supply hub i and refueling station j, between supply hub i and liquefaction plant l, and between liquefaction plant l and refueling station j, (f_{ij},f_{il},f_{lj}).

The objective function is formulated to maximize profit of producing LNG, delivering LNG, and selling it at each LNG refueling station. The profit is defined as the annual revenue from the sale of LNG less the annual cost of producing the LNG fuel.

Revenue is represented by the number of LNG gallons sold at the current market price of LNG sold per gallon, which is derived from the retail diesel price at a specific station location, (RP_j). Each prospective station can only generate profit if it is active which is indicated by the binary variable, (X_{jt}). Since truck demand at each station location is given in diesel trucks, we introduce the parameter ,(δ), to represent the percentage of trucks that will refuel with LNG fuel. In order to derive the fuel demand (D_{jr}) from truck demand, we multiply the number of LNG trucks by (α), which is the parameter for LNG full-tank size.

Facility costs are handled as discrete points (e.g. station sizes) instead of along a continuous interval. The cost of procuring natural gas from the market hub is identified by parameter, (SP_i). The station fixed cost (SFX_{jT}) at station location j and technology T is actually defined as a variable, not as a parameter for reason that will be discussed later. Supporting the calculation of, (SFX_{jT}) are parameters, (SFF, B, SI, LI) representing standard station fixed cost, base station fixed cost, standard station storage cost, and large station storage cost. Since conventional and modular LNG have different components, conventional technology has a variable operation and maintenance cost, ($SVOM_t$) dependent on station volume, while modular LNG technology has a fixed operation and maintenance cost, ($SFOM_{t*}$). However, both technologies have a variable electricity cost depending on LNG volume, (SE_{jT}). Modular LNG stations have an additional fixed capital cost in the case, where a pipeline must be constructed or extended to reach the station, (PCC_j). The cost to transport natural gas from supply hubs to the modular LNG technology station is represented by (PT_{ij}). Since conventional technology requires LNG to be produced at LNG liquefaction plant, the cost of building the LNG liquefaction plant is represented by (PF_{ls}) at plant location l with capacity s. For conventional technology, natural gas must first be transported by pipeline to the plant with cost, (PT_{il}) and then it needs to be transported by truck to the station with cost, (TT_{lj}).

The objective function is constrained by a combination of both the physical limitations of facilities and realistic assumptions about the LNG industry, all represented inside this mathematical formula. The reason we use a derived LNG price is that natural gas wholesale spot markets have over the past few years decoupled from oil markets (Hartley et al., 2008). By generating prices based on costs to deliver, this captures the variation that might take place in markets over time and across space.

Objective:

$$Max\ Profit = \sum_{ij}(RP_j \cdot f_{ij}) + \sum_{lj}(RP_j \cdot f_{lj}) - Cost \tag{1}$$

where

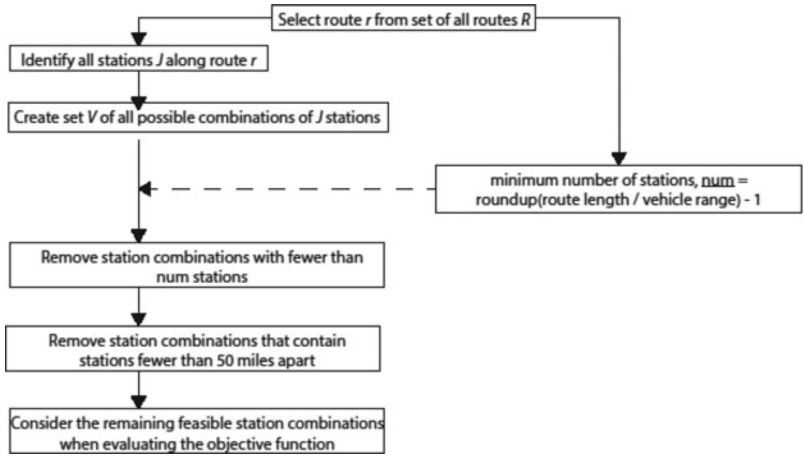
$$\begin{aligned} Cost = {} & supply\ cost + station\ fixed\ capital\ cost \\ & + variable\ station\ cost\ (operation\ \&\ maintenance\ \&\ electricity) \\ & + pipeline\ transport\ cost\ from\ supply\ to\ station \\ & + LNG\ liquefaction\ plant\ capital\ cost\ (includes\ fixed\ operation\ and\ maintenance) \\ & + variable\ LNG\ liquefaction\ plant\ cost\ (electricity) \\ & + pipeline\ transport\ cost\ from\ supply\ to\ LNG\ plant \\ & + truck\ transport\ cost\ from\ LNG\ liquefaction\ plant\ to\ station \end{aligned}$$

$$\begin{aligned} & \sum_i (SP_i \cdot (\sum_j f_{ij} + \sum_l f_{il})) + \sum_{jt} SFX_{jt} + \sum_{lj}((SVOM_t + SE_{jt}) \cdot f_{lj}) + \sum_{ij}(SE_{jt*} \cdot f_{ij}) \\ & \quad + \sum_{ij}(PT_{ij} \cdot f_{ij}) + \sum_{ls}(Y_{ls} \cdot PF_{ls}) + \sum_{il}(PE_l \cdot f_{il}) + \sum_{il}(PT_{il} \cdot f_{il}) + \sum_{lj}(TT_{lj} \cdot f_{lj}) \end{aligned} \tag{2}$$

The maximization of the objective function is subject to numerous constraints. These constraints include the feasibility constraint such stations will not be constructed in such a way that the distance between any two stations is greater than the vehicles' range, the profitability constraint that ensures that the sum of station profits along each route must be non-negative, as well as technical constraints to ensure demand is satisfied. The feasibility constraint is explained in Appendix B. Briefly, we employ an algorithm, illustrated in Figure 2, to identify all possible combinations of station construction along a particular route and remove combinations that cannot be traversed given a particular vehicle range. To speed computation we also remove station combinations where stations are fewer than 50 miles apart.

The remaining constraints are explained in detail in Appendix C. The first set of constraints is necessary in defining the total LNG station cost. The reason why fixed station cost is a variable and not a parameter is because a single station location actually represents a collection of stations,

Figure 2: Feasibility Algorithm



which we can assume in real world to be several adjacent stations. This collection of adjacent stations represented, as a single location in the model must be a variable because it needs to be able to adapt to the demand at that location. It is formulated this way is because we found that even at the largest station sizes we provided, the station capacity was still insufficient to satisfy the demand in later years. This complicates the model's representation of a single station too because it needs to be able to represent total costs, total size/capacity, and total flow as if it were multiple stations stacked on top of each other. In order to do this, we introduce new variables and parameters specifically for this purpose. (NS_{jT}), represents the number of the largest capacity stations, which in our case is ($CM = 60{,}000$ LNG gallons/day) for conventional station technology (t) or ($BM = 10{,}000$ LNG gallons/day) for modular LNG technology (t^*) at location (j). (RC_j) represents the remaining demand that still needs to be satisfied, in excess of the demand already satisfied by the set of full-capacity stations at location (j), which only applies to conventional stations because modular LNG is only offered at a single size. Finally, (SFX_{jT}) can be calculated from the previous two variables along with the station flow.

The first series of constraints (C.1–C.10) define the number of full-capacity stations (60,000 LNG gallons for conventional and 10,000 LNG gallons for modular LNG) required to meet the demand of LNG, with constraints (C.1–C2) representing both technologies, (C.3–C.6) representing modular LNG technology (t^*) and constraints (C.7–C.10) representing conventional technology (t). The different technologies require separate constraints. They express the mathematical connection between the number of stations built or needed and the actual LNG flow coming into a station. Constraints C.3, C.4 and C.7, C.8 together acts like an upper bound for the number of stations that can be built.

The next three constraints (C.11–C.13) apply only to the conventional station technology (t). Due to the fact that conventional technology is offered at multiple sizes (s), station sizing is more flexible which also requires more constraints to capture this flexibility. Previously, we discussed how the model decides the number of full-capacity stations (NS_{jT}), but for example, if there is some leftover volume that might not require building another full capacity 60,000 LNG gallon/day station, the model may choose to build smaller sizes, that is multiple standard stations (15,000 LNG gallons/day) not exceeding 60,000 LNG gallons/day. The purpose of these constraints is to define first the remaining capacity (RC_j) at station (j) and then define the number of remaining stations (RN_j) required, with each additional station size equaling 15,000 LNG gallon/day of additional storage.

Constraints C.14–C.16 require that the total station fixed cost at a site must be greater than the annual individual station fixed cost times the number of stations at the site. Constraints C.17 and C.18 are technology and size configurations constraints that require that at most one technology may be chosen per site and at most one LNG plant size may be chosen per site at time of construction. Constraints C.19–C.21 are facility capacity constraints for plant and supply field. Constraints C.22–C.24 require that the quantity of fuel dispensed at each site must not exceed demand. Constraints C.25–C.28 require that if a station combination is selected to activate a route, all stations belonging to the station combination must be constructed. Constraints C.29 to C.32 are intertemporal relationships that require stations constructed in previous years to remain built in later years and that flows in later years must be greater than or equal to flows in previous years. Lastly, Constraints C.33 and C.34 require that all decision variables be either one or zero and that all flow quantities and facility counts be non-negative.

3. RESULTS AND DISCUSSION

To date, despite the strongest market for commercial truck sales in almost a decade and a historic gap between low natural gas prices and high oil prices, America's natural gas highway is struggling to take hold (Tita, 2014). Our analysis confirms this trend and finds that only certain regional markets have sufficient traffic density in combination with higher diesel prices compared to the U.S. national average to give investors a sufficient return on capital to incentivize station construction without government intervention.

Our analysis shows that despite the fuel cost advantages that might result from some limited regional natural gas transportation network buildouts, the development of a U.S. national natural gas transportation network will be encumbered by high initial investment costs for new cross country infrastructure relative to the fully discounted, incumbent oil-based network. Rather, we find that a concentrated regional focus in key markets for early investment is the least-cost strategy to initiate the development of natural gas transportation networks in the United States.

We find that the level of profitability of natural gas fueling infrastructure is more highly correlated with access to a high volume of traffic flows of freight movements than with the locus of surplus supplies of natural gas. Thus, initiatives to introduce natural gas freight fueling businesses in regions with stranded or inexpensive gas resources (natural gas supplies that lack sufficient demand to be commercialized) run a greater risk of failure than efforts to introduce natural gas fueling infrastructure along major freight routes in California, the Great Lakes region and the U.S. Mid-Atlantic.

Figure 3 shows the concentration of trucking traffic on U.S. interstates with thickness of line representing those routes with the heaviest truck traffic flows as per the U.S. Department of Transportation Freight Analysis Framework. As the figure shows, the West Coast and the Great Lakes region are among the heaviest flows in the United States and therefore may have the highest potential for a new fuel.

Our regional analysis, under 0.2 percent initial market penetration of LNG line-haul, shown in Figure 4, shows that California and the U.S. Great Lakes/Northeast regions, which have a relatively high level of truck traffic, have the greatest commercial potential at present and could play a key role in the network development. In the case of LNG heavy-duty trucking networks, California, which has a concentrated highway routing for long-haul trucking delivering goods from ports, is uniquely positioned to launch a profitable natural gas network. The costs to provide ded-

Figure 3: Concentration of Truck Traffic

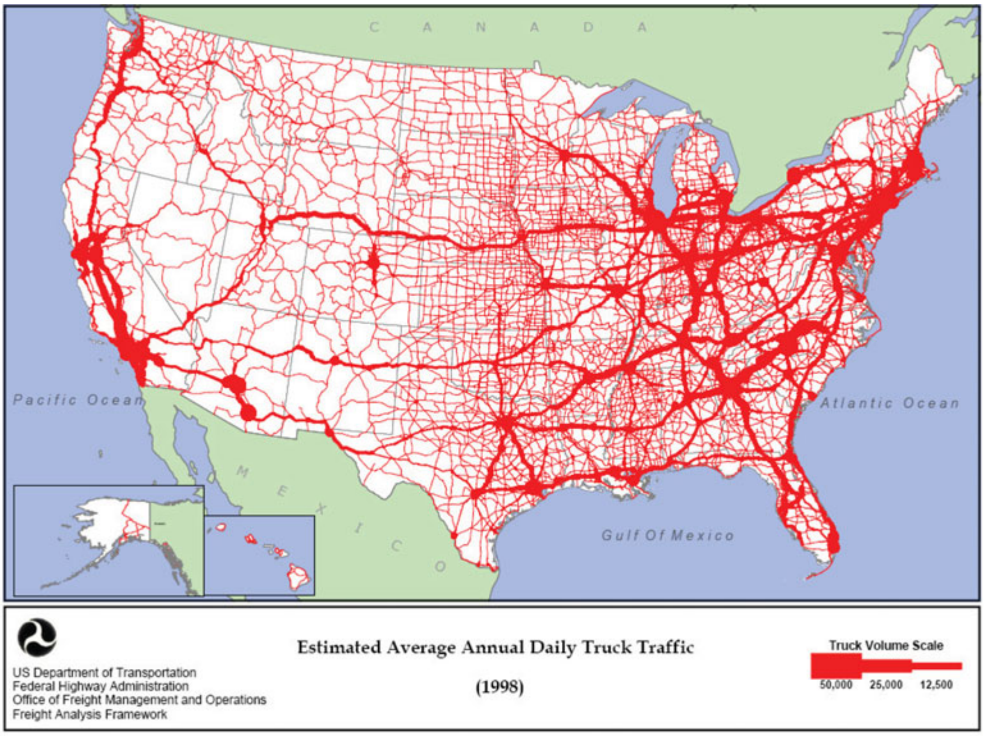


Figure 4: LNG station build out results under 0.2%, No Subsidy, Year 2030

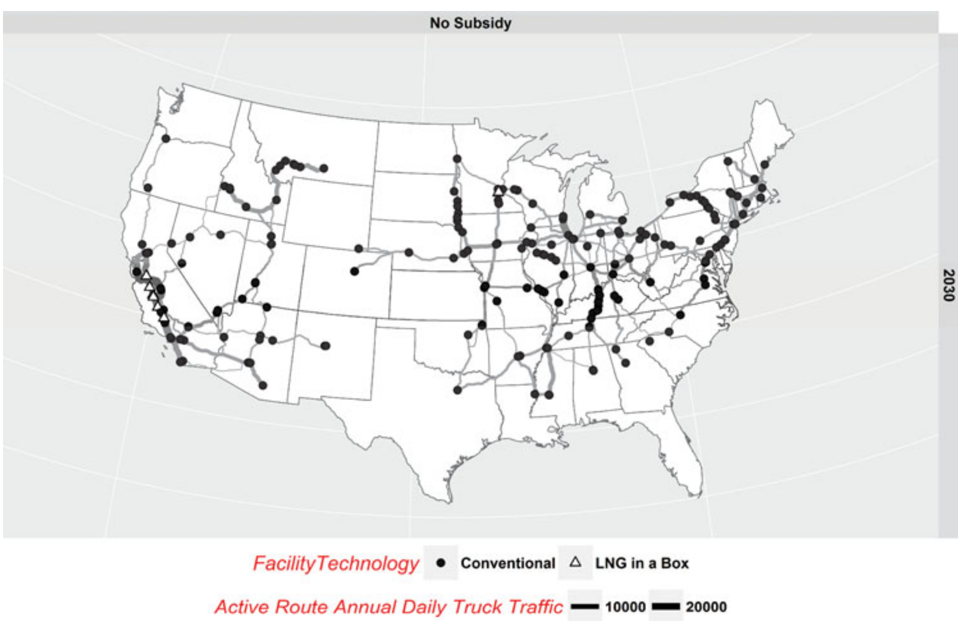
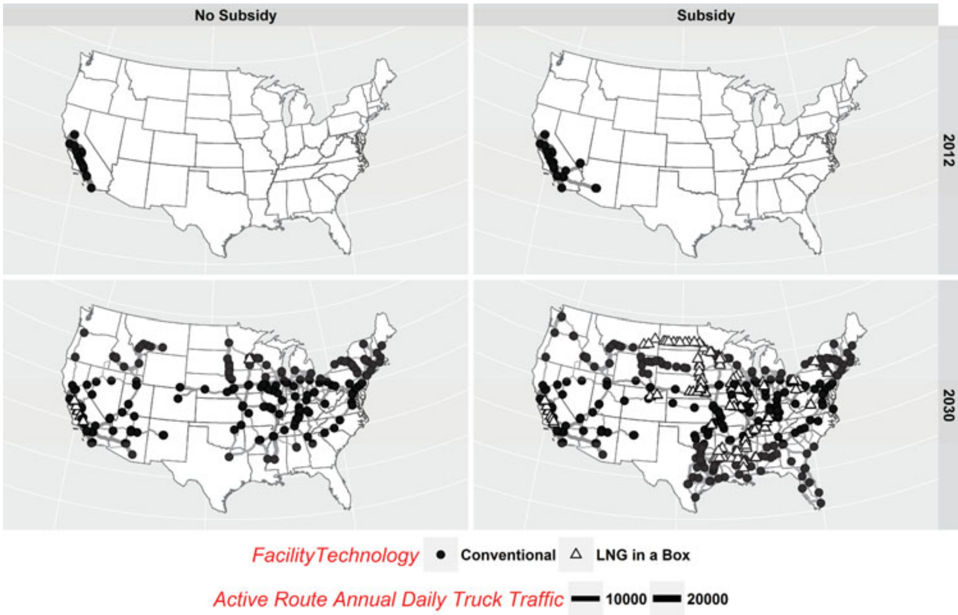


Figure 5: Dynamic LNG station build out results under 0.2%



icated coverage for LNG across California are estimated to be less than $100 million. The Great Lakes and mid-Atlantic areas are also well-positioned to incubate a natural gas transportation network.

Figure 5 shows the spatial distribution of stations over time under a market with a penetration rate of 0.2 percent, roughly double today's penetration rate. Under no subsidy case, initial station development in 2012 looks relatively familiar to the subsidy scenarios with development localized to mainly to California. By 2030, the model seeds development in a new hotspot region in the Mid-West and Mid-Atlantic. Not surprisingly, profitable station build-out patterns favor the regions with higher traffic volumes which characterizes locations like California, Wisconsin, Illinois, Kansas City, Nashville, Cincinnati, New York, and Boston areas as well as areas with higher diesel prices like California, Illinois, and New York. Conventional technology is highly favored over modular LNG early on. This is mainly due to the high upfront cost of the modular LNG technology.

Figure 5 also suggests some very interesting station technology implications. We have already seen that at higher fuel delivery volumes, modular LNG has the potential to save the station operator on transportation costs because at higher volumes of demand, trucking fuel from LNG liquefaction plants becomes increasingly expensive. As the national network grows and as technology costs drop, modular LNG also becomes an important technology for connecting remote regions in the U.S. Mid-Continent (Heartland & Mountain regions) with coverage gaps to the larger network of stations since traffic volumes are generally lower and therefore unsuitable for larger liquefaction infrastructure systems. Although conventional technology still dominates the network, modular LNG is notably more widespread in the subsidy scenario, suggesting that the higher cost of that technology is a barrier and lowering its costs with a subsidy would support its deployment. These scenarios show that policy choices could influence the competition between LNG station technologies.

Figure 6: Dynamic LNG station price difference per diesel gallon equivalent (DGE) under 0.2% market penetration

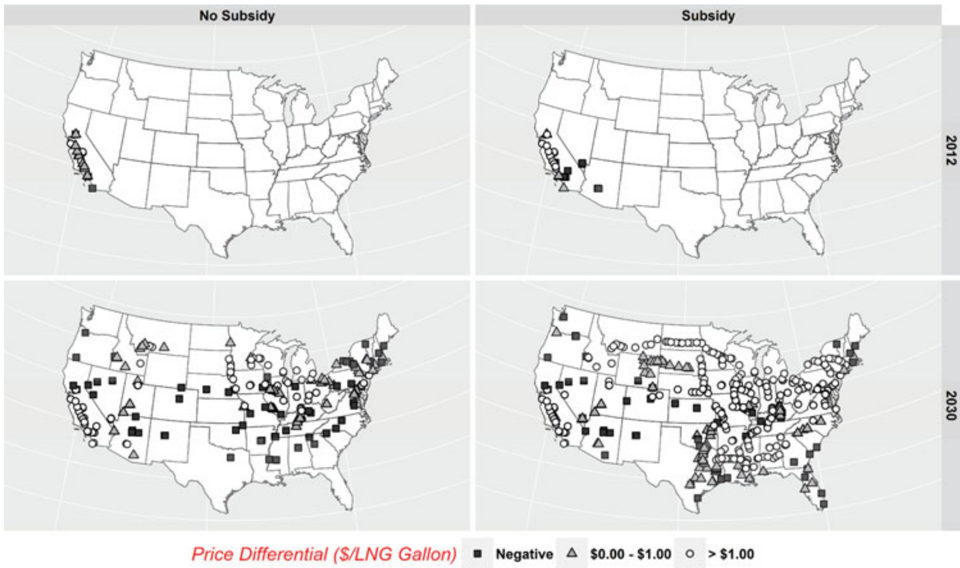


Figure 6 shows a combination of potentially profitable stations in white circles and light-grey triangle and shows unprofitable stations in dark grey squares. Assuming a 0.2 percent market penetration and under no subsidy base case, the map illustrates many competitive station price differentials greater than a dollar per dge exist in California and Midwest. Interestingly, high volume routes in Texas and Florida lag behind other regions as early adopters. In the case of Texas, low diesel prices may be a contributing factor for slow network growth. Florida lacks a high volume route into the state, despite a large flow of traffic exiting from its ports. If the U.S. could somehow spur LNG truck demand to double, a significant portion of truck routes could potentially be covered by 2030 without the help of a subsidy. However, a competitive national network of LNG stations (white circle) is unlikely to spawn without the help of subsidies or higher diesel prices.

Figure 7 provides a good illustration of the hub and spoke like evolution of the LNG network, which begins in California and eventually expands into the Mid-West and the East Coast. By 2030, the network of truck routes in the no subsidy scenario augments to the point where West almost meets East. In the subsidy case, the networks connect by 2030. It is evident that the evolution of the network is very similar in both scenarios; the only difference is the specific timing of when certain routes are selected for deployment.

To test the sensitivity of the profitability of a national network to the number of trucks on the road, we analyze a scenario where double the current number of trucks would be operating with LNG fuel. We compare our modeling results against four case study scenarios: 1) a 50% subsidy to station costs (or the equivalent of a 50% cost breakthrough) under a 0.1 percent penetration rate; 2) a 50% subsidy to station costs (or the equivalent of a 50% cost breakthrough) under a 0.2 percent penetration rate; 3) no subsidy under a 0.1 percent penetration rate; 4) no subsidy under a 0.2 percent penetration rate. Table 1 summarizes our results.

We find the initial number of natural gas vehicles on the road at the start of infrastructure investment has a dramatic impact on the development of the natural gas refueling network. For

Figure 7: Trucking route deployment by scenario across the modeling horizon under 0.2% market penetration

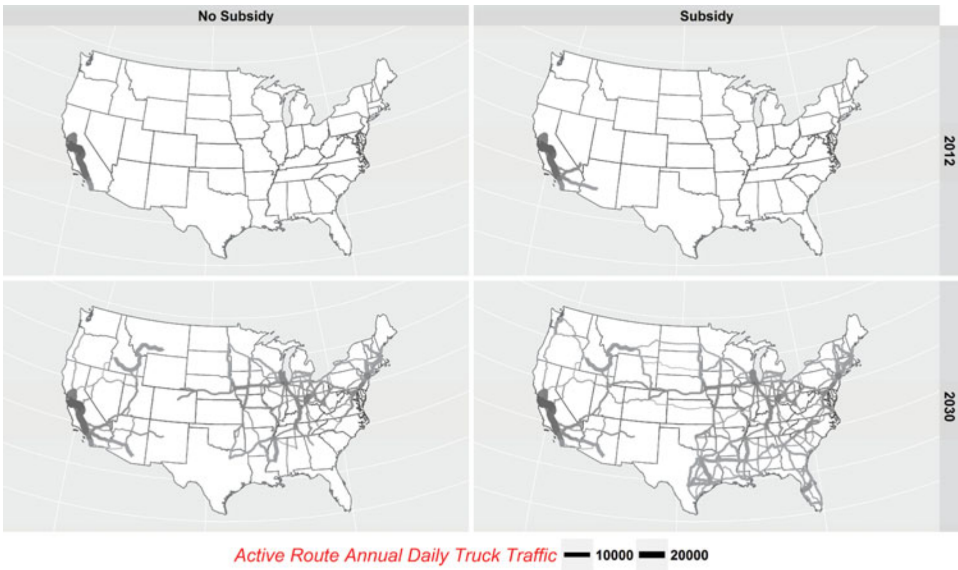


Table 1: Summary of Sensitivity Analysis Results

	0.1% Initial Penetration Rate			0.2% Initial Penetration Rate		
	Summary	**Route Completion 2015**	**Route Completion 2030**	**Summary**	**Route Completion 2015**	**Route Completion 2030**
No Subsidy	Network only builds in California.	0%	2%	Network begins in California and extends Eastward. Northeast and Great Lakes regions begin construction in 2025.	3%	55%
50% Subsidy	Network begins in California and extends eastward.	2%	6%	Network begins construction in California, Arizona and Nevada. Construction in the Great Lakes begins in 2015.	28%	76%

instance, in the no subsidy scenario, a doubling of the number of natural gas long-haul truck in operation at the outset of the model (0.2% penetration rate) results in 55% of the interstate highway network being refuel-able with natural gas by 2030. This compares to the no subsidy scenario with a penetration rate of 0.1% which yields only 2% network coverage by 2030. Again, with penetration rate of 0.1%, the total network coverage by 2030 reaches only 6% even under a 50% station capital subsidy or the high-diesel price scenario. However, with a 0.2% penetration rate, the network coverage by 2030 climbs to 76% under the 50% subsidy scenario.

4. CONCLUSIONS AND IMPLICATIONS FOR POLICY

The deeply entrenched incumbency of oil-based fuels and their well-established infrastructure distribution provide a formidable barrier to the transition to alternative fuels. Even for a fuel such as LNG, which enjoyed a deep cost discount to diesel from 2011 to 2013, establishing a competitive fueling network will be challenging. Moving LNG into the long-haul trucking fleet could prove the most pliable of the options for fuel-switching based on commercial factors. That is because the turnover rate for Class 8 vehicles is fairly rapid compared to other kinds of vehicle stocks (three years, for example, compared to 10 to 14 years for light-duty vehicles) and vehicle ownership tends to be concentrated in large corporate fleets whose vehicles have high miles utilization per year and who can scale up more quickly than individual vehicle owners to shift vehicle technologies. Our analysis also suggests that efforts to focus preliminary alternative fuels development on high volume, heavily trafficked roadways first would ensure the highest likelihood of success. We find that the level of profitability of alternative fueling infrastructure, in the case of natural gas, is more highly correlated with access to a high volume of traffic flows of freight movements than with the locus of surplus fuel feedstock supplies. This is in contrast to studies on biofuels where high costs of transport make the location of feedstock more directly relevant to the economics of development.

Large fleet owners will not be willing to make investments in alternative fuel vehicles unless they are assured of dedicated fueling station availability for their entire travel route. Thus, our scenario analysis suggests that the best way to promote an alternative fuel, such as LNG, into the heavy-duty trucking sector would be to focus initially on the highest volume freight routes such as California and the upper Midwest and then eventually commercial factors will encourage investment to branch out to other hotspot regions such as the Mid-Atlantic.

Conceptually, focusing on a handful of large fleets that could commit to substantial purchases of LNG trucks in a particular regional market makes commercial sense and is consistent with the current commercial climate. For example, UPS ordered about 700 natural gas tractors in 2013 alone, showing the viability of getting adoption of the additional trucks via a fleets purchasing model.

Our analysis would support new efforts by the state of California and the U.S. Federal government to explore sustainable freight policies that might enable alternative fuels. Our findings suggest that previous work on hydrogen that projects a cluster based strategy for fueling station infrastructure could also work for national level freight corridors, if coordinated with the freight industry, which is increasingly operating on a hub and spoke paradigm.

ACKNOWLEDGMENTS

The authors would like to thank Colin Carter, Kevin Novan and Joan Ogden of UC Davis and Brandon Owen and Mike Farina of GE for their counsel and advice on this study. We also thank GE Ecomagination and the Sloan Foundation for their generous support to our program. Funding for this work included a grant from the California Energy Commission which does not in any way endorse or advocate the positions set forth in this research.

REFERENCES

Agnolucci, P. and W. McDowall (2013). "Designing future hydrogen infrastructure: Insights from analysis at different spatial scales." *International Journal of. Hydrogen Energy* 38: 5181–5191. https://doi.org/10.1016/j.ijhydene.2013.02.042.

Almansoori, A., and N. Shah (2009). "Design and operation of a future hydrogen supply chain: Multi-period model." *International Journal of Hydrogen Energy* 34: 7883–7897. https://doi.org/10.1016/j.ijhydene.2009.07.109.

Alternative Fuels Data Center, (2017a). Alternative Fueling Station Counts by State [WWW Document]. URL http://www.afdc.energy.gov/fuels/stations_counts.html (accessed 2.11.17).

Alternative Fuels Data Center, (2017b). Natural Gas Vehicles [WWW Document]. URL http://www.afdc.energy.gov/vehicles/natural_gas.html (accessed 2.11.17).

Apanel, G., and E. Johnson (2004). "Direct methanol fuel cells–ready to go commercial?" *Fuel Cells Bulletin.* 2004: 12–17. https://doi.org/10.1016/S1464-2859(04)00410-9.

Balinski, M.L. (1965). "Integer Programming: Methods, Uses, Computations." *Management Science*. 12: 253–313. https://doi.org/10.1287/mnsc.12.3.253.

Bandivadekar, A., C. Evans, T. Groode, J. Heywood, E. Kasseris, M. Kromer, and M. Weiss (2008). On the Road in 2035: Reducing Transportation's Petroleum Consumption and GHG Emissions. 2008. Massachusetts Institute of Technology.

Charles S. and R.W.S. Revelle (1970). "Central Facilities Location," in *Geographical Analysis*. pp. 30–42.

Church, R., and C. Revelle (1972). The Maximal Covering Location Problem. Pap. Reg. Sci. Assoc. 6.

Dagdougui, H. (2012). "Models, methods and approaches for the planning and design of the future hydrogen supply chain." *International Journal of Hydrogen Energy* 37: 5318–5327. https://doi.org/10.1016/j.ijhydene.2011.08.041.

Dominguez-Faus, R., C. Folberth, J. Liu, A.M. Jaffe, and P.J.J Alvarez (2013). "Climate change would increase the water intensity of irrigated corn ethanol." *Environmental Science Technololgy* 47: 6030–6037. https://doi.org/10.1021/es400435n.

Freeman, L.C., (1977). "A Set of Measures of Centrality Based on Betweenness A Set of Measures of Centrality Based on Betweenness," in: *Sociometry*. American Sociological Association, pp. 35–41.

Fulton, L., and M. Miller (2015). Strategies for Transitioning to Low-Carbon Emission Trucks in the United States, A white paper from the Sustainable Transportation Energy Pathways program at UC Davis and the National Center for Sustainable Transportation.

Greene, D.L., P.N. Leiby, B. James, J. Perez, M. Melendez, A. Milbrandt, S. Unnasch, D. Rutherford, and M. Hooks (2008). Analysis of the transition to hydrogen fuel cell vehicles and the potential hydrogen energy infrastructure requirements. Oak Ridge National Laboratory (ORNL).

Grønhaug, R., and M. Christiansen (2009). Supply Chain Optimization for the Liquefied Natural Gas Business, in: Al., L.B. et (Ed.), *Innovations in Distribution Logistics*, Lecture Notes in Economics and Mathematical Systems. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-92944-4_10.

Guo, Z., J. Deride and Y. Fan (2016). "Infrastructure planning for fast charging stations in a competitive market." T*ransportation Research. Part C Emerging Technology* 68: 215–227. https://doi.org/10.1016/j.trc.2016.04.010.

Hakimi, S. (2014). Optimum Locations of Switching Centers and the Absolute Centers and Medians. *Informs* 12: 450–459.

Hamacher, Z.D.H.W. (2002). Facility Location—Applications and Theory.

Hartley, P.R., K. B. Medlock, and J.E. Rosthal (2008). "The Relationship of Natural Gas to Oil Prices." *The Energy Journal* 29(3): 47–65. https://doi.org/10.5547/ISSN0195-6574-EJ-Vol29-No3-3.

Hugo, A, P. Rutter, S. Pistikopoulos, A. Amorelli, and G. Zoia (2005). Hydrogen infrastructure strategic planning using multi-objective optimization. *International Journal of Hydrogen Energy* 30, 1523–1534. https://doi.org/10.1016/j.ijhydene.2005.04.017.

Johnson, N., C. Yang, and J. Ogden (2008). "A GIS-based assessment of coal-based hydrogen infrastructure deployment in the state of Ohio." *International Journal of Hydrogen Energy* 33: 5287–5303. https://doi.org/10.1016/j.ijhydene.2008.06.069.

Kelley, S., and M. Kuby (2013). On the way or around the corner? Observed refueling choices of alternative-fuel drivers in Southern California. *Journal of Transportation Geography* 33: 258–267. https://doi.org/10.1016/j.jtrangeo.2013.08.008.

Knittel, C.R., (2012). Reducing petroleum consumption from transportation. *Journal of Economic Perspectives* 26: 93–118. https://doi.org/10.3386/w17724.

Kristoffersen, T.K., and S.E. Fleten (2010). Stochastic Programming Models for Short-Term Power Generation Scheduling and Bidding, in: Bjørndal, E., Bjørndal, M., Pardalos, P.M., Rönnqvist, M. (Eds.), *Energy, Natural Resources and Environmental Economics, Energy Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 187–200. https://doi.org/10.1007/978-3-642-12067-1.

Krupnick, A.J., (2011). Will Natural Gas Vehicles Be in Our Future? Resources for the Future. Issue Brief. 11-06.

Kuby, M., and S. Lim (2007). "Location of alternative-fuel stations using the Flow-Refueling Location Model and dispersion of candidate sites on Arcs. Networks." *Spatial Economics* 7: 129–152. https://doi.org/10.1007/s11067-006-9003-6.

Kuby, M., and S. Lim (2005). The flow-refueling location problem for alternative-fuel vehicles. Socioecon. Plann. Sci. 39, 125–145. https://doi.org/10.1016/j.seps.2004.03.001.

Kuby, M., L. Lines, R. Schultz, Z. Xie, J.-G. Kim, and S. Lim (2009). "Optimization of hydrogen stations in Florida using the Flow-Refueling Location Model." *International Journal of Hydrogen Energy* 34: 6045–6064. https://doi.org/10.1016/j.ijhydene.2009.05.050.

Lee, A., N. Parker, R. Dominguez-Faus, D. Scheitrum, Y. Fan, A.M. Jaffe (2015). Sequential Buildup of Liquefied Natural Gas Refueling Infrastructure System for Heavy-Duty Freight Trucks. Proc. Present. to Transp. Res. Board Annu. Meet.

Lin, Z., C. Chen, J. Ogden, and Y. Fan (2008a). "The least-cost hydrogen for Southern California." *International Journal of Hydrogen Energy* 33: 3009–3014. https://doi.org/10.1016/j.ijhydene.2008.01.039.

Lin, Z., J. Ogden, Y. Fan, and C. Chen (2008b). "The fuel-travel-back approach to hydrogen station siting." *International Journal of Hydrogen Energy* 33: 3096–3101. https://doi.org/10.1016/j.ijhydene.2008.01.040.

Medlock, K.B., (2012). "Modeling the implications of expanded US shale gas production." *Energy Strategy Review* 1: 33–41. https://doi.org/10.1016/j.esr.2011.12.002.

Medlock, K.B.I.I.I. and Peter Hartley (2006). *The Baker Institute World Gas Trade Model*. Cambridge: Cambridge University Press.

Melo, M.T., S. Nickel, and F. Saldanha-da-Gama (2009). "Facility location and supply chain management—A review." *European Journal of Operations Research* 196: 401–412. https://doi.org/10.1016/j.ejor.2008.05.007.

Morrison, G.M., N.C. Parker, J. Witcover, L.M. Fulton, and Y. Pei (2014). "Comparison of supply and demand constraints on US biofuel expansion." *Energy Strategy Review* 5: 42–47. https://doi.org/10.1016/j.esr.2014.09.001.

Murthy Konda, N.V.S.N., N. Shah, and N.P. Brandon (2011). "Optimal transition towards a large-scale hydrogen infrastructure for the transport sector: The case for the Netherlands." *International Journal of Hydrogen Energy* 36: 4619–4635. https://doi.org/10.1016/j.ijhydene.2011.01.104.

Nakićenović, N., (1987). The automobile road to technological change?: diffusion of the automobile as a process of technological substitution. Internat. Inst. for Applied Systems Analysis.

National Petroleum Council (2012). Advancing Technology for America's Transportation Future: Fuel and Vehicle Systems Analyses: Natural Gas Analysis.

National Renewable Energy Laboratory (2015). 2015 Natural Gas Vehicle Research Roadmap (Draft).

Ni, J., N. Johnson, P. Manager, J.M. Ogden, A.C. Yang, J. Johnson, T. Studies, C. Davis, A. Hydrogen (2005). Estimating Hydrogen Demand Distribution Using Geographic Information Systems (GIS). Institute of Transportation Studies.

Ogden, J., and M. Nicholas (2011). "Analysis of a 'cluster' strategy for introducing hydrogen vehicles in Southern California." *Energy Policy* 39: 1923–1938. https://doi.org/10.1016/j.enpol.2011.01.005.

Ogden, J.M., M.M. Steinbugler, and T.G. Kreutz (1999). "A comparison of hydrogen, methanol and gasoline as fuels for fuel cell vehicles: implications for vehicle design and infrastructure development." *Journal of Power Sources* 79: 143–168. https://doi.org/10.1016/S0378-7753(99)00057-9.

Parker, N., Y. Fan, Y., and J. Ogden (2010). "From waste to hydrogen: An optimal design of energy production and distribution network." *Transportation Research, Part E Logistics Transportation Review* 46: 534–545. https://doi.org/10.1016/j.tre.2009.04.002.

Parker, N.C., J.M. Ogden and Y. Fan (2008). "The role of biomass in California's hydrogen economy." *Energy Policy* 36: 3925–3939. https://doi.org/10.1016/j.enpol.2008.06.037.

Rosenberg, E., A. Fidje, K. Aamodt, C. Stiller, A. Mari, S. Møller-holst (2010). "Market penetration analysis of hydrogen vehicles in Norwegian passenger transport towards 2050." *International Journal of Hydrogen Energy* 35: 7267–7279. https://doi.org/10.1016/j.ijhydene.2010.04.153.

Sevtsuk, A., and M. Mekonnen (2012). Urban Network Analysis. SimAUD '12 Proceedings of the 2012 Symposium on Simulation for Architecture and Urban Design. Article No. 18 .

Sharpe, B.R. (2013). Examining the costs and benefits of technology pathways for reducing fuel use and emissions from on-road heavy-duty vehicles in California. University of California, Davis. Research Report UCD-ITS-RR-13-17.

Simchi-Levi, D., P. Kaminsky and E. Simchi-Levi (2003). *Managing the Supply Chain: The Definitive Guide for the Business Professional*. McGraw Hill Professional.

Strachan, N., N. Balta-ozkan, D. Joffe, K. Mcgeevor and N. Hughes (2009). "Soft-linking energy systems and GIS models to investigate spatial hydrogen infrastructure development in a low-carbon UK energy system." *International Journal of Hydrogen Energy* 34: 642–657. https://doi.org/10.1016/j.ijhydene.2008.10.083.

Straight, B. (2014). Alt fuels: Beyond natural gas | Running Green content from Fleet Owner [WWW Document]. FleetOwner. URL http://fleetowner.com/running-green/alt-fuels-beyond-natural-gas (accessed 2.11.17).

TIAX for America's Natural Gas Alliance (2010). Liquefied Natural Gas Infrastructure [WWW Document]. Am. Gas Assoc. Tech. Rep. URL https://www.aga.org/tiax-natural-gas-vehicle-market-analysis (accessed 2.13.17).

Tita, B., (2014). "Slow Going for Natural-Gas Powered Trucks" [WWW Document]. *Wall Street. Journal.* URL https://www.wsj.com/articles/natural-gas-trucks-struggle-to-gain-traction-1408995745 (accessed 2.13.17).

Toregas, Constantine, R. Swain, C. ReVelle and L. Bergman (1970). "The Location of Emergency Service Facilities." *Operations Research* 19(6): 1363–1373. https://doi.org/10.1287/opre.19.6.1363.

Wang, Y.-W., and C.-C. Lin (2009). "Locating road-vehicle refueling stations." *Transportation Research, Part E Logist. Transportation Review* 45: 821–829. https://doi.org/10.1016/j.tre.2009.03.002.

Werpy, M.R., D. Santini, A. Burnham, and M. Mintz (2010). Natural gas vehicles: Status, barriers, and opportunities. Argonne National Laboratory. ANL/ESD/10-4 Energy Systems Division.

Zeng, W., I. Castillo, and M.J. Hodgson (2010). "Aggregating Data for the Flow-Intercepting Location Model: A Geographic Information System, Optimization, and Heuristic Framework." *Geographical Analysis* 42: 301–322. https://doi.org/10.1111/j.1538-4632.2010.00795.x.

APPENDIX A: TRUCK DEMAND MODEL

The trucking demand model outputs the percent of truck miles traveled that are economic to fuel with LNG (the penetration rate) as a function of the price gap between LNG and diesel fuel.

To calculate the volume of trucks converting to LNG, the truck fleet turnover uses historical data on the distribution of the truck fleet by model year in 2012, the survival of trucks as they age and a sales-to-scrap ratio to grow the fleet over time.

For each year, a set percentage of trucks of a given age are scrapped based on historical survival rate of trucks. Historically 50% of trucks are scrapped by the time they are 16 years old and very few trucks survive to age 30.[14] The fleet grows if the sales/scrap ratio is greater than one. For 2012–2020 the ratio is assumed to be 1.1 and for 2021–2030 it is assumed to be 1.2 (Sharpe, 2013). These assumptions increase the size of the truck fleet from 3.2 million in 2012 to 3.8 million trucks in 2030.

Older trucks stay in operation but they shift to lower mileage applications so their contribution to the energy demand (% of truck miles traveled) is less than their population suggests. To capture the distribution of truck miles by age and use, we use the Vehicle Inventory and Use Survey from 2002. The 2002 data is the most recently available data available to assess truck miles between trucks of different classifications.

We are able to use our calculation of the distribution of use to determine the percent of truck miles traveled by trucks in each age group, which can then be applied in the fleet turn over model to determine the fraction of truck miles that are traveled by trucks in each model year. This allows us to track the influence of new trucks over time on the potential LNG market share.

The decision to purchase LNG trucks instead of diesel trucks is represented in the model by a discounted 3-year payback rule. If LNG trucks offer a 3-year or less payback then the LNG truck is purchased; otherwise if the payback is longer than 3 years or non-existant, a diesel truck is purchased. This decision is based on the overall pattern of the purchasing decision for new vehicles in the heavy duty sector where initial owners tend to hold vehicles for three to four years before reselling. The payback is sensitive to the cost of the LNG truck, annual mileage of the truck, the relative fuel economy, and the maintenance costs differential reported in Table A.1. We do not specify the maintenance costs or secondary market resale value of the vehicle, which is among the commercial factors in fleet decision-making to invest in LNG trucks. While we acknowledge that these factors are a constraint in the early stages of a new network, we assume that these factors would sort themselves out as the LNG network takes hold, if the payback for a shift to LNG were sufficiently attractive. The model finds the price gap between diesel and LNG that makes LNG the better deal so while diesel price is important it is not a parameter in the model. The broad distribution

14. Emissions Factor (EMFAC) 2007 model

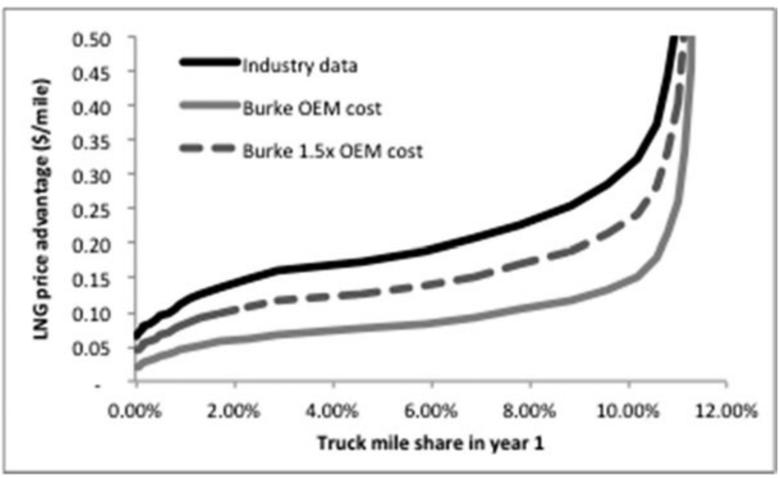
of annual travel for new trucks found in the VIUS 2002 data leads to a demand curve for LNG trucks holding all other parameters constant shown in Figure A.1. At small price gaps only the truck with the highest travel will find LNG competitive but as the price gap grows, larger percentages of the new truck buyers will find LNG attractive until the single year market is saturated.

Table A.1: Input Data

	Industry data[a]	Burke OEM cost	Burke 1.5X OEM cost
Incremental LNG truck cost	$35,000	$18,000	$27,000
Incremental O&M cost* ($/mile)	$0.0276	$0.0276	$0.0276
LNG Fuel economy (miles/DGE)	5.1	5	5
Diesel Fuel economy (miles/DGE)	6	6.3	6.3
Required rate of return	0.12	0.12	0.12

* assumed to be 30% higher for LNG
[a] Data from Westport corporate analysis division, ITS engine models

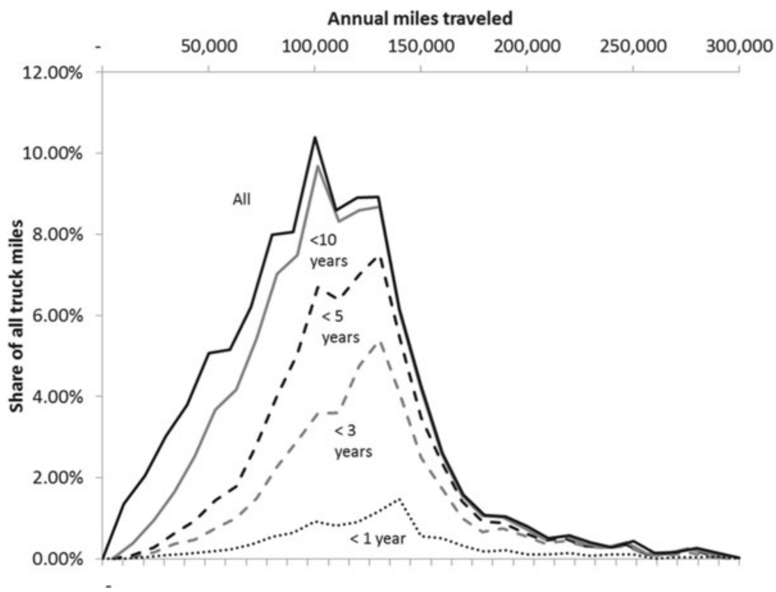
Figure A.1: LNG Price Discount versus LNG Truck VMT Share



The market for new heavy-duty trucks in the coming years will be substantial. Between 2014 and 2025 roughly 2.7 million new trucks will be purchased or 76% of the total fleet in 2025, creating a ready market for natural gas vehicles, if commercial incentives are evident. While the average truck travels 80,000 miles a year and remains in service for roughly 16 years, we analyze the movements for the largest trucks that travel more extensively than this 80,000 miles per year average and thereby would have the greatest payback for a shift to natural gas.

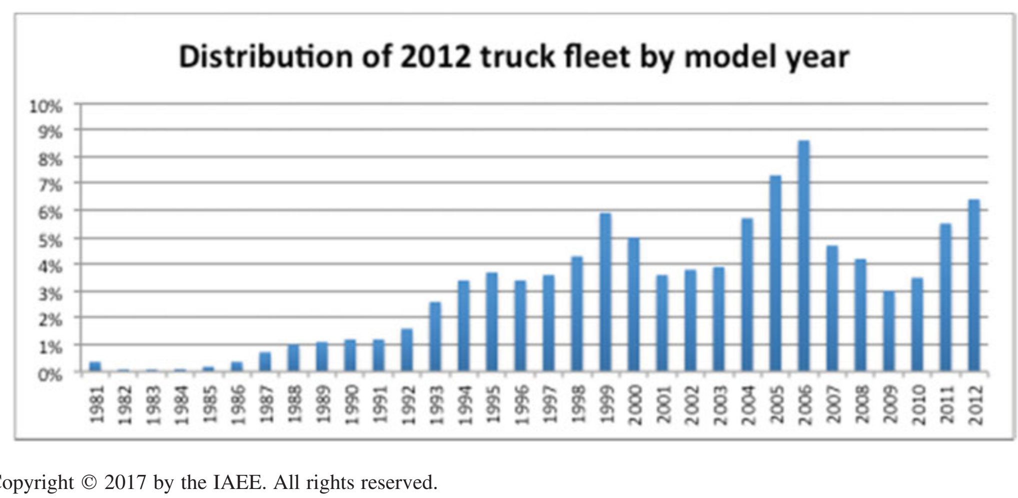
Figure A.2 below shows the percentage of total truck miles travelled by different populations of trucks. We are most interested in trucks that will travel more than 125,000 miles a year because these would have the shortest payback period for a shift to natural gas fuel. As the figures show there are a limited number of trucks that travel over 175,000 miles per year and these are predominantly trucks less than 3 years in age. The average age is low because typically, large fleets, which represent a significant share of the heavy duty market, replace their vehicles about every three years. Looking by age, data shows that new trucks (one year old) travelling around 125,000 miles per year are responsible for around 1.75% of all truck miles. Trucks travelling between 100,000 and 110,000 miles per year are responsible for just over 10% of the all truck miles in the United States.

Figure A.2: Truck VMT Distribution by Age of Vehicle



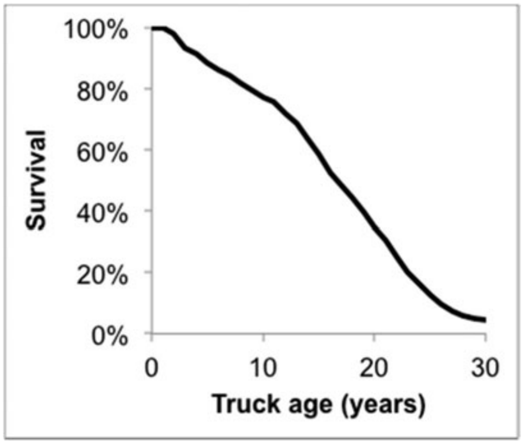
The current distribution of the truck fleet by model year is displayed in Figure A.3. The starting distribution of truck ages is not uniform due to a number of factors including air quality regulations and economic cycles. For instance, the 2005 and 2006 model years are disproportional large due to impending air quality regulations on MY2007 trucks. The economic downturn depressed sales in 2009. Model years 2011 and 2012 represent a rebound in the market as the economic recovered and the deferred purchases were finally made. This distribution of truck ages informs the number of new truck purchases that will be made in any given year that could potentially shift to natural gas vehicles.

Figure A.3: Distribution of 2012 Truck Fleet by Model Year



The following figure, Figure A.4, shows the percentage of truck fleet survival by number of years of service. Retirement is derived for a given year by calculating the loss of trucks in the next year and dividing by the trucks remaining in that year.

Figure A.4: Truck Survival Rate by Age



The model of course is a simplified version of real world demand to switch to LNG fuel based solely on the economic potential of LNG demand and therefore represents the maximum possible demand among truck operators. In reality, other considerations can weigh large, such as availability of trained mechanics, expected performance and vehicle failure rates, risk assessment to future prices etc. The simulation at this stage provides for no dampening functions on the uptake of LNG to account for experience with LNG (training maintenance staff, expected failure rates, etc.), the availability of truck models/engine sizes or other factors beyond incremental capital, incremental O&M and fuel savings.

APPENDIX B: FEASIBLE STATION COMBINATIONS

We create a flow-based algorithm similar to the one created by Kuby and Lim (2005) that will generate all feasible combinations of refueling stations able to refuel a specific origin-destination truck route. Figure B.1 below is taken from (Kuby and Lim, 2005) illustrates a simple network example with links and nodes and will be used to explain the objective. The numbers represent distance in arbitrary units. Take for example, there is a truck at origin O and it is trying to travel to destination D. The two nodes A and B represent existing potential refueling stations. Since we assumed that trucks could be refueled at their origin and destination, we have simplified the problem to ignore the round-trip. Thus, there are only three unique sets of station combinations that can be built and they are {A}, {B}, or both {A, B}. But can all of these potential station combinations refuel the given route? This depends on the vehicle range, which we assume to be roughly 300–400 miles based on current LNG vehicle technology. If the vehicle range is greater than 500 miles, any location can refuel the trip or rather no station is even required, assuming stations at both origin and destination already exists. If we set vehicle range anywhere between 300 and 500 miles, trucks would require at least one station and that one station must be placed at node B because A would not provide sufficient range for the rest of the trip. However, both A and B is also an option if say there is another O-D route that intersects at node A. Finally, if the vehicle range is less than 300 miles, then no single node can refuel the trip if locations are restricted to those nodes.

A number of relevant extensions to the original FCLM exist but perhaps, the most relevant extension for siting potential LNG stations for heavy-duty vehicles has to do with the distance between stations because LNG trucks are range limited. Given that LNG fuel contains roughly 60% of the energy density of diesel fuel and storage of LNG is very expensive, more strategic planning must be conducted to locate refueling stations. Kuby and Lim (2005) extended the original flow-capturing model to take into account vehicle range. Instead of a single station being able to satisfy flow, they considered a flow to be captured or refueled only if an adequate combination of station(s) were allocated appropriately in order to refuel the vehicle along the path.

The task of optimally selecting a combination of stations is another problem in itself because there could be any number of possible station combinations or infinite that could refuel any given route. Kuby and Lim (2005) developed an algorithm, which could enumerate all possible vertex node combinations along a line segment and return only those feasible node sets, which were capable of refueling the vehicle's range on a hypothetical network. Since the previous method led to sub-optimality, Kuby and Lim (2007) further developed an algorithm that would identify possible candidate nodes on arcs, in between vertices. While these methods are suitable for siting an entirely new facility concept where there is no equivalent existing infrastructure, they are not necessary for siting alternative fueling stations because using existing infrastructure is easier and perhaps more relevant in the near term at least.

Flow-based approaches have been applied to solve various problems in transportation. Wang and Lin (2009) developed a flow-based set-covering model for locating refueling stations using a mixed integer programming method based on vehicle-routing logistics. Their work presents a case study for intercity highways in Taiwan that minimizes the total costs of locating refueling stations while ensuring that every vehicle has sufficient fuel to complete all origin-destination routes, parameterized by vehicle range. However, this article has a number of less realistic simplifications including equal weight on passing flows and the specific model formulation limits it to solving moderately sized problems. Kuby et al. (2009) applied their FCLM to optimizing the locations of hydrogen stations in Florida using real world traffic and demographic data. They proposed a mixed integer linear programming (MILP) model that selects refueling stations, which maximize the number of trips (or vehicle miles traveled) on the transportation network, parameterized by vehicle range and number of stations that must be built. Although the authors present a convincing analysis of optimal facility location and real-world refueling station rollout, such analysis would be incomplete if they didn't account for system fixed capital costs and variable costs.

All above-mentioned flow-based models have been classified by Agnolucci and MacDowall (2013) as "local scale" models, which determine optimal locations of refueling stations in a relatively small geographical space. These models are likely unsuitable to solve real-life, large-scale problems because they share common characteristics namely, a single period planning horizon, a single facility type, size, and cost, a single product, and a single location-allocation decision. Although they are clearly insufficient to cope with many realistic scenarios, they provide an important backbone for many of the problems we are interested in solving. Therefore, literature suggests that many practical applications from supply chain management (SCM) can be incorporated into and used to extend the basic facility location model framework.

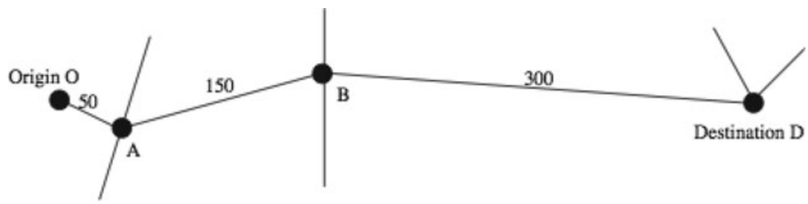
Supply chain management (SCM) is the practical addition to the FCLM because in addition to generic facility location setup, it considers many real world applications such as procurement, production, inventory, distribution, and routing (Melo et al., 2009). With a purpose of solving regional scale problems, SCM is the process of planning the operations of the supply chain in an efficient manner in an attempt to achieve some objective such as minimizing system costs or maximizing total system profits (Simchi-Levi et al., 2003). In addition to answering questions of where, it answers questions of the number, capacity, size, technology, flow, and cost of all facilities belonging to the logistics network. Therefore, supply chain modeling with economic cost and spatial analysis has been a popular method for evaluating new vehicle technology rollout from a systems point of view.

Hugo et al. (2005) presented a generic model for optimal long-term planning and design of future hydrogen supply chains. They utilized a mixed integer optimization program to provide optimal integrated investment strategies across a wide range of supply chain decision-making stages. Although the study was comprehensive in its analysis of the supply chain, it was only applied to a small geographic region case study. Moreover, demand for hydrogen is uniformly represented across the geographic region and the long-term demand trajectory is assumed to follow the popular s-shaped curve, which has been used to represent the diffusion patterns of new energy technology rollouts including vehicles (Nakićenović, 1987). Almansoori and Shah (2009) proposed a more detailed multi-period model for optimizing the operation of the future hydrogen supply chain. This model tries to address the space and time variation of demand during the evolution of the hydrogen station network over a long-term planning horizon. However, their demand treatment is still limited because their model setup assumes coarse spatial grid blocks and the familiar s-shaped curve. Konda et al. (Murthy Konda et al., 2011) presented a similar detailed methodology but applied it to a large-scale real case study for the Netherlands. In contrast to previous studies, the authors explicitly addressed the spatial aspects by utilizing statistical information on geography and demographics to identify regional demand variation. Still, their treatment of long-term temporal demand is inadequate given that they only applied a linear growth projection on Dutch passenger car fleet populations in order to estimated fuel-cell vehicle market share.

An algorithm is implemented and executed in Python programming language to generate a master list of feasible station combinations for each O-D route after eliminating supersets. This algorithm will determine whether each combination can refuel the given route and remove ones from the list if it can't, assuming a vehicle range. The linear reference tool in ArcGIS is used to produce mile markers for every station along every route employing the Contiguous USA Equidistant Conic projection to ensure distances are accurately measured. For each route, the algorithm iterates through all station combinations, moving from one station to the next station in the set to determine whether that combination is feasible. For this algorithm, we assume that existing LNG refueling stations are already present at both the origins and destinations so there is no need to investigate the round trip. We also assume that once a hypothetical truck reaches a station, it is able to refuel up to the full-specified vehicle range. If the hypothetical truck is not able to reach the next station in the combination set, it will be eliminated and the algorithm continues with the next station combination. Once the algorithm is finished, it outputs the recorded set of feasible LNG station combinations.

The computational burden of this algorithm are immense, especially because it has to generate $\sum_{k=1}^{n} C_k^n$ combinations for each route, where n is the number of stations and k is the required number of stations to refuel that route. However, this would be inefficient if we just searched every possible combination, instead we can just use C_k^n by first calculating the required stations needed to refuel a route by simply dividing the total route distance by the vehicle range. Next, assess whether the number is a whole number, if it is, subtract one, else round down and the new number should give the minimum necessary stations to refuel that given route. For example, if the route distance is 100 miles and the vehicle range is 50 miles, it would require a single station to refuel it but now if the range was 60 miles, it would still only require one station. Even with this heuristic, the national model could produce upwards of 100,000 different route combinations, which would be too cumbersome for the optimization model to solve in reasonable time. In effort to reduce the number of combinations even more, the following simplifying heuristics are introduced. The basic premise is to again reduce the station redundancy because if a truck stops at one station, it would not be so necessary for the truck to stop again 50 miles down the road again, especially if the goal of the LNG industry is to try to expand network coverage.

Figure B.1: Flow Based Model Illustration



From the list of all feasible station combinations, they must adhere to the guiding assumptions for generating a valid combination for the model. These assumptions were verified by industry experts in major oil companies, commercial refueling station operators, and energy corporations heavily investing in LNG. Most of these assumptions have already been implicitly mentioned but for explicit understanding, they are as follows:

(1) Trucks travel from a single origin to a single destination without route deviation.
(2) Trucks will travel on their shortest path with respect to shortest time.
(3) Truck refueling stations must be located less than a mile from a truck's shortest route.
(4) Trucks begin at their origin with a full tank of fuel and can refuel at their destination, thereby guaranteeing a round trip.
(5) Both Origin and Destination (Cities) already have built, existing LNG refueling stations.
(6) Origin (i) Destination (j) pairs cannot be circular tours; they must start at point A and end at point B.

Origin Destination *distance* matrices are assumed to be symmetrical meaning that ordered pairs i, j and j, i are identical and thus can be treated as an unordered pair.

APPENDIX C: OPTIMIZATION CONSTRAINTS

Constraints C.1 and C.2 apply to new stations (nj) only because they can be of both technologies. Constraint C.1 states that the total number of full capacity stations ($\sum NS_{jT}$) and standard capacity stations must be greater than or equal to the binary build variable ($\sum X_{jT}$) at that station for all new stations (nj). This ensures that if a station of any technology is built, there must be at least one type of station there. Constraint C.2 states that the total number of full capacity stations ($\sum NS_{jT}$) must be less than the total flow ($\sum_i f_{ij} + \sum_l f_{lj}$) for all new stations ($nj$). This constraint ensures that if either technology flows are zero, then there must be no station present.

Constraints C.3 and C.7 state that the number of full capacity stations by respective technologies (NS_{jt*}/NS_{jt}) must be less than or equal to the sum of all respective LNG flow entering the station ($\sum_i f_{ij}/\sum_l f_{lj}$), iterating over each station. Constraints C.4 and C.8 state that the number of full capacity stations by respective technologies (NS_{jt*}/NS_{jt}) must be less than or equal to some large number (M), multiplied by the binary decision variable to build each respective station technology (X_{jt*}/X_{jt}), iterated over each station. These constraints are important because if either flow or build decisions are zero, then the number of stations must be zero too. Constraint C.5, C.6 and C.9, C.10 do the opposite of C.3, C.4 and C.7, C.8 because they act like a lower bound for the number of stations. Constraints C.5 and C.9 state that the number of full capacity stations by respective technologies (NS_{jt*}/NS_{jt}) and remaining stations (RN_j) (in case of conventional technology) must be greater than or equal to zero if the respective station technology is built (X_{jt*}/X_{jt}), iterated over each station.

Finally, constraints C.4 and C.8 are the most important in the set because they actually define the number of full-capacity stations by respective technologies (NS_{jt*}/NS_{jt}) that are required to satisfy the station's demand flow, whereas the previous constraints were more like logical constraints. It does this by stating that number of full capacity stations must be greater than or equal to the total respective flow ($\sum_i f_{ij}/\sum_i f_{ij}$) divided by full respective station technology capacity (BM/CM) as a function of technology type for each station. Since (NS_{jt*}/NS_{jt}) is required to be integer, this expression should automatically round up to add another station.

$$\sum_T NS_{jT} + RN_j \geq \sum_T X_{jT} \quad \forall\ nj \tag{C.1}$$

$$\sum NS_{jT} \leq \sum_i f_{ij} + \sum_l f_{lj} \quad \forall\ nj \tag{C.2}$$

$$NS_{jt*} \leq \sum_i f_{ij} \quad \forall\ nj \tag{C.3}$$

$$NS_{jt*} \leq M \cdot X_{jt*} \quad \forall\ nj \tag{C.4}$$

$$NS_{jt*} \geq X_{jt*} \quad \forall\ nj \tag{C.5}$$

$$NS_{jt*} \geq \frac{\sum_i f_{ij}}{BM} \quad \forall\ nj \tag{C.6}$$

$$NS_{jt} \leq \sum_l f_{lj} \quad \forall\ j \tag{C.7}$$

$$NS_{jt} \leq M \cdot X_{jt} \quad \forall\ j \tag{C.8}$$

$$NS_{jt} + RN_j \geq X_{jt} \quad \forall\ j \tag{C.9}$$

$$NS_{jt} \geq \frac{\sum_i f_{ij}}{CM} - 1 \quad \forall\ j \tag{C.10}$$

Constraint C.11 defines the remaining station capacity (RC_j) to be greater or equal to the sum of all conventional flows ($\sum_l f_{lj}$) going into the station minus the capacity (NS_{jt}) already taken care of by the full-capacity stations, iterated for every conventional station ($j \in t$). Constraint C.12

is a capacity constraint on the remaining capacity (RC_j), which states that it cannot surpass the maximum capacity (CM) of a single full capacity built conventional station (X_{jt}), iterated for all conventional stations ($j \in t$). This means that if build decision is zero, then the remaining capacity will be zero as well. Finally, constraint C.13 is important because it defines the remaining number of stations (RN_j) to be greater than or equal to the remaining capacity (RC_j) divided by the capacity of a single standard station ($I = 15{,}000$ LNG gallons/day), which is the size that stations can be upgraded, iterated for all conventional stations ($j \in t$).

$$RC_j \geq \sum_l f_{lj} - NS_{jt} \cdot CM \quad \forall \; j \in t \tag{C.11}$$

$$RC_j \leq CM \cdot X_{jt} \quad \forall \; j \in t \tag{C.12}$$

$$RN_j \geq \frac{RC_j}{I} \quad \forall \; j \in t \tag{C.13}$$

These first 13 constraints were really sub-constraints that lead up to the definition of the total station fixed cost, (SFX_{jT}) which we define next. Constraint C.14 applies only to modular LNG stations ($j \in t^*$) and it states that the total station fixed cost must be greater than or equal to the number of modular LNG stations (NS_{jt^*}) multiplied by the annualized (CRF) station cost (SF_{t^*}) (including pipeline connection cost (PCC_j) if necessary) which is then multiplied by the technology-learning rate (STR) and subsidy rate (Z), and lastly, plus the fixed operational/maintenance cost ($SFOM_{t^*}$). Constraints C.15 and C.16 both define the total station fixed cost for conventional stations ($j \in t$). The costs in constraint C.15 can be broken down to three parts: 1) cost of storage for full-capacity stations at 60,000 LNG gallons/day, ($NS_{jt} \cdot LI \cdot CRF$), 2) cost of storage for remaining station capacity at 15,000 LNG gallons/day, ($RN_j \cdot SI \cdot CRF$), and 3) cost of basic station components, ($X_{jt} \cdot B \cdot CRF$) (everything excluding storage). Each cost component is multiplied against each respective binary decision variable and then the sum of all costs is multiplied by technology learning rate (STR) and station subsidy rate (Z). All constraints should be iterated over all stations based on technology. Constraint C.16 is a special constraint in the case of already existing conventional stations (EXT) or (FAF)/city stations ($ej \in t$). In this case, we assume that some percentage (E) of the original standard station fixed cost ($B + SI$) has already been financed and it is applied towards the total annualized (CRF) station cost (SFX_{jt}), iterated over all existing conventional stations.

$$SFX_{jt^*} \geq NS_{jt^*} \cdot ((SF_{t^*} + PCC_j) \cdot CRF \cdot (1 - STR^{Year-2012}) \cdot (1 - Z) + SFOM_{t^*}) \quad \forall \; j \in t^* \tag{C.14}$$

$$SFX_{jt} \geq ((NS_{jt} \cdot LI \cdot CRF) + (RN_j \cdot SI \cdot CRF) + (X_{jt} \cdot B \cdot CRF)) \cdot (1 - STR^{Year-2012}) \cdot (1 - Z) \quad \forall \; j \in t \tag{C.15}$$

$$SFX_{jt} \geq SFX_{jt} - (B + SI) \cdot E \cdot CRF \quad \forall \; ej \in t \tag{C.16}$$

The next set of constraints is characterized as single station configuration constraints. This applies to both liquefaction plants and LNG stations. Constraint C.17 states that the binary station build variable, (X_{jT}) summed over both station technologies (T) must be less than or equal to one, iterated over new station (nj) locations because we assume for simplicity that it is invalid to have two technologies at one station. Similarly, constraint C.18 states that the plant binary build variable, (Y_{ls}) summed over all plant sizes (s) must be less than or equal to one, iterated for every plant location (l), for the same reason as previously.

$$\sum_T X_{jT} \leq 1 \quad \forall \; nj \tag{C.17}$$

$$\sum_s Y_{ls} \leq 1 \quad \forall \; l \tag{C.18}$$

Constraints C.19–C.21 represent facility capacity constraints for plant and supply locations. Constraint C.19 states that for each plant location l, all natural gas volumes by pipeline ($\sum_i f_{il}$) being distributed to the plant must be less than or equal to the build plant's (Y_{ls}) capacity (M_{ls}) for a given size (s). Constraint C.20 states that for each supply location i, all natural gas by pipeline ($\sum_j f_{ij} + \sum_l f_{il}$) leaving the supply hub (i) must be less than or equal to the supply (S_i) available. Notice that there is no equivalent capacity constraint for LNG stations and the reason is because this constraint is already implicitly stated within the previous series of constraints in the number of stations and station capacities. As previously mentioned, truck and pipeline capacity is assumed negligible as a constraint in the model because the industry sees no immediate concern for lack of truck or pipeline capacity given current LNG demand forecasts. Constraint C.21 represents the conservation of flow between natural gas flow entering ($\sum_i f_{il}$) a plant location and LNG leaving ($\sum_j f_{lj}$) the plant, for each plant (l).

$$\sum_i f_{il} \leq \sum_s Y_{ls} \cdot M_{ls} \quad \forall \; l \tag{C.19}$$

$$\sum_j f_{ij} + \sum_l f_{il} \leq S_i \quad \forall \; i \tag{C.20}$$

$$\sum_i f_{il} = \sum_j f_{lj} \quad \forall \; l \tag{C.21}$$

The next set of constraints (C.22–C.24) is important because they are the reason why demand is satisfied. Constraint C.22 states that flow received by pipeline ($\sum_i f_{ij}$) for each new modular LNG station ($nj \in t^*$) must be less than or equal to the sum over for all built modular LNG stations' (X_{jt^*}) demand (D_{jr}) multiplied by (F_{Year}) for the list of routes (r_j) that intersect station ($nj \in t^*$). (F_{Year}) is a special parameter that inflates annual truck demand in order to anticipate for future expansion as would happen in the real world. Constraint C.23 is similar except it applies to all conventional stations. Finally, constraint C.24 states that all flows ($\sum_i f_{ij} + \sum_l f_{lj}$) going into either station technologies must be less than or equal to that total demand (D_{jr}) for routes that are built (R_r) within the list of routes (r_j) for all stations (j).

$$\sum_i f_{ij} \leq \sum_{r_j} (X_{jt^*} \cdot F_{Year} \cdot D_{jr}) \quad \forall \; nj \in t^* \tag{C.22}$$

$$\sum_l f_{lj} \leq \sum_{r_j} (X_{jt} \cdot F_{Year} \cdot D_{jr}) \quad \forall \; j \in t \tag{C.23}$$

$$\sum_i f_{ij} + \sum_l f_{lj} \leq \sum_{r_j} (R_r \cdot F_{Year} \cdot D_{jr}) \quad \forall \; j \tag{C.24}$$

Constraints C.25, C.26, C.27 collectively represent the flow-based sub-model. Constraint C.25 pertains to truck routes ($r \in yr$), which require at least one new station in order to make it refuelable, and it states that there must be at least one station combination (h), which makes (V_h) equal to 1, that belongs to route (r) that can successfully refill O-D route (W_r). Constraint C.26 pertains to trucks routes ($r \in nr$), which don't require a new station, and it states that all EXT/FAF conventional stations ej that belong to the route must be greater than or equal to the binary route build decision (W_r). Basically if a particular route is selected, then every EXT and FAF station on the route must be built. Constraint C.27 states that for a given route ($r \in yr$) that needs a station, if a station combination (V_h) is open, then every new station ($nj \in j_{rh}$) within the combination (h) and route (r) must be built/open ($\sum_T X_{jT}$), in order for (v_h) to equal 1; it is zero otherwise. Constraint C.28 is the same as the previous, except it is for existing EXT/FAF conventional stations that need to be built (X_{jt}) if the station combination (V_h) is open.

$$\sum_h v_h \geq W_r \quad \forall \ r \in yr \,|\, h \in h_r \tag{C.25}$$

$$X_{jt} \geq W_r \quad \forall \ r \in nr \,|\, \mathrm{e}j \tag{C.26}$$

$$\sum_T X_{jT} \geq V_h \quad \forall \ r \in yr; \quad \forall \ h \in h_r; \quad \forall \ \mathrm{n}j \in j_{rh} \tag{C.27}$$

$$X_{jt} \geq V_h \quad \forall \ r \in yr; \quad \forall \ h \in h_r; \quad \forall \ \mathrm{e}j \in j_{rh} \tag{C.28}$$

Then, finally, we have a set of constraints to make a connection between results from one year to the next year. Constraint C.29, C.30, and C.31 state that all routes ($\sum_{r^*} W_r$), LNG liquefaction plants ($\sum_{l^*s} Y_{ls}$), and LNG stations ($\sum_{j^*T} X_{jT}$) built in the previous year must also be built in all future years where the superscript (*), represents recorded facilities (r^*, l^*,j^*) built in the previous years. To simplify the computation, we relaxed the part of the constraint that forces specific technologies or sizes to be built. Constraint C.32 states that the sum of all new flows ($\sum_l f_{il}$) to LNG liquefaction plant in future years must be greater than all previous year's flows ($\sum_{l^*} f_{il}$).

$$\sum_{r^*} W_r \geq 1 \quad \forall \ r^* \tag{C.29}$$

$$\sum_{l^*s} Y_{ls} \geq 1 \quad \forall \ l^* \tag{C.30}$$

$$\sum_{j^*T} X_{jT} \geq 1 \quad \forall \ j^* \tag{C.31}$$

$$\sum_l f_{il} \geq \sum_{l^*} f_{il} \quad \forall \ l^* \tag{C.32}$$

Finally, Constraint C.33 is a binary constraint on the above decision variables meaning their objective values must be either zero or one. Constraint C.34 collectively constrains all flows and station counts within the supply chain model to be zero or positive.

$$X_{jt}; \ Y_{ls}; \ V_h; \ R_r \in (0,1) \tag{C.33}$$

$$f_{ij}, \ f_{lj}, \ f_{il}, \ NS_{jT}, \ RN_j - Non.Negativity \tag{C.34}$$