



Optimizing Fueling Decisions for Locomotives in Railroad Networks

V. Prem Kumar*

Michel Bierlaire*

06 October 2011

Report TRANSP-OR 111006

Transport and Mobility Laboratory (TRANSP-OR)

School of Architecture, Civil and Environmental Engineering (ENAC)

École Polytechnique Fédérale de Lausanne (EPFL)

`transp-or.epfl.ch`

Abstract

Even though rail transportation is one of the most fuel efficient forms of surface transportation, fueling costs are one of the highest operating cost head for railroad companies. In US, unlike Europe, fueling costs are indeed, by far, the single highest operating cost. For larger companies with several thousands of miles of rail network, the fuel bills often run into several billions of dollars annually. The railroad fueling problem considered in this paper has three distinct cost components. Fueling stations charge a location dependent price for the fuel in addition to a fixed contracting fee over the entire planning horizon. In addition, the railroad company must also bear incidental and notional costs for each fuelling stop. It is imperative that the number of fueling stops between an origin and destination should be restricted to avoid unnecessary delays. This paper proposes a mixed integer linear program model that determines the optimal strategy for contracting and fuel purchase schedule decisions that minimizes overall costs under certain reasonable assumptions. This model is tested on a large, real-life problem instances. This mathematical model is further enhanced by introducing several feasible MIP cuts. This paper compares the efficiency of different MIP cuts in order to reduce the run-time. Lastly, the paper concludes with an observation that even though the problem scale was expected to diminish the model performance, it was indeed noted that run-time and memory requirements are fairly reasonable. It thus establishes that this problem must be looked beyond the prism of heuristics and other approximate algorithms for actual implementation at railroad companies.

Keywords: Scheduling, Large-scale Optimization, Railroads, Fueling Decisions

1. Introduction

About three-fourths of the world's railroads operate with diesel fuel. Even though most of the railroad network in Europe is electrified, US railroads depend almost entirely on diesel fuel. Association of American Railroads has calculated that it takes roughly seven gallons of fuel to move one ton of ordinary freight from one coast to the other in USA¹. This is only about one-fourth of the fuel required to transport the freight by roads. But it is equally important to note that fueling costs have been increasing over the last few years. According to a report published by *Oak Ridge National Laboratory*, fueling costs at Amtrak (Passenger and Freight Rail Company) have gone up from 6% to 11% of their entire budget between 2004 and 2008. This is in spite of the fact that railroads continue to be most fuel efficient form of bulk transport compared to surface transport by 17% and air transport by 33%.

¹ Transcript of the statement by Phillip Longman before US House of Representatives Transportation Infrastructure Committee on Jan 28, 2009 (source: <http://www.millennium-institute.org/resources/elibrary/papers/>)

Fuel selling prices vary diversely between locations due to differences in local taxes, distribution costs, marketing costs and other factors. According to INFORMS RAS Problem Solving Competition (2010), as of August 1, 2006, one gallon of diesel costs one of the Class-I railroad company \$2.2057 in Atlanta, GA, but \$2.2823 at Augusta, GA. One of the major challenges faced by a railroad is to determine a fuelling strategy for its entire locomotive fleet so that costs are kept at a minimum. Fueling costs for a railway network usually have two components – fixed and variable. The fixed cost relates to providing the necessary infrastructure for fueling at the selected station yards while the variable cost is a factor of the fueling stops and the amount of fuel consumed.

We now review some literature for the fuel purchasing problem as observed in transportation modes other than railroads. The problem of fueling cost optimization has been first studied in the context of airline industry. Stroup and Wollmer (1991) minimize the total fuel cost for an airline flight schedule, subject to aircraft capacity and other side constraints, and formulate and solve the problem as a LP. Zouein et al. (2002) consider the set of all scheduled flights on known routes to minimize the total fuel purchasing costs, and formulate and solve the problem as a variant of the multi-period inventory planning problem. There are many papers in the last five years that have developed and implemented models to optimally locate fuel stations so that flow in a network is maximized. Kuby and Lim (2005) and Kuby et al. (2005) presented a mixed-integer program to find the optimal location of refueling facilities for a new form of alternative fuel and solve the same using a greedy algorithm framework. Kuby and Lim (2007) extended their model by adding candidate facilities along network arcs. Khuller et al. (2007) studied a series of fueling schedule problems to find the optimal travel route that minimize fuel costs needed to travel from an origin to a destination, or to visit a set of predetermined points. This paper assumes that fueling locations are known in advance. Upchurch et al (2009) extended the fuel station location model by Kuby and Lim (2007) to account for fuel stations with capacity limitations. Wang and Lin (2009) proposed a flow-based set covering model for road network with an electric vehicle to locate refueling (recharging) stations that would minimize the total facility cost while ensuring that vehicle can never run out of fuel during a journey. This model assumes that each vehicle is refueled every time it passes a refueling station, and the formulated model is solved by a commercial solver (with an embedded branch-and-bound algorithm). No consideration was given to potentially varying fuel prices at different stations, the delay penalty associated with fueling stops, and hence the potential benefit from strategically scheduling fueling activities (e.g., it may be optimal for a vehicle to skip refueling at some of the fuel stations). Lim and Kuby (2010) presented a heuristic algorithm that solve for the optimal refueling station locations to maximize the flow such that there are restrictions on refueling with a given number of facilities.

This problem of optimizing fuel purchase decisions in railroads is not an extensively researched topic. Most of the problems in the domain of optimizing fueling costs in railroads considered in the literature have their own cost dimensions. The problem considered by Nourbakhsh and Ouyang (2010) is

the closest to the problem defined in this work, even though there are some subtle differences in the nature of the problem. Firstly, the distance between successive station yards in our problem is assumed to be such that locomotive would never run out of fuel during its journey. Thus there is no provision for emergency fueling between station yards in our problem. The second difference arise from the fact that this paper assumes a restriction on the number of locomotives fueled at a particular station yard for a time period while we assume a restriction on the volume of fuel filled for a particular time period (day) at a particular station yard within the planning horizon in our problem. Nourbakhsh and Ouyang (2010) formulate the problem as a mixed integer program (MIP) and then decompose it into two sub-problems to be solved using Lagrangean relaxation. The easier problem that involves the selection of minimal yards for the entire network is solved greedily while the difficult sub problem minimizes the number of fueling points and cost of fuel for every locomotive and is solved using a polynomial-time shortest path algorithm through a certain proposition. We will prove in section 4 that the proposition used in Nourbakhsh and Ouyang (2010) would not be applicable for our problem with an instance and thus would suggest better methods for solution.

Nag and Murthy (2010) suggest a greedy algorithm for the locomotive refueling problem and claim that this method would be more appropriate compared to formulating and solving as MIP for two reasons – one being that solving a MIP model would require a commercial solver and the other being that large problem instances of real-life would produce better results through a heuristic, rather than an exact method. However, railway companies across the world are usually large companies with high operational costs due to the nature of business and high investment costs. In addition, there are many other functions in railways operations, such as locomotive assignment, rolling stock management, platform scheduling etc. that would require optimization algorithms. So it is unlikely that a railway company would cut corners to save money on commercial solvers. Optimization solvers are used extensively for several decades at all major top airline and even railway companies. Moreover, freeware solvers such as COIN-OR and GLPK provide efficient algorithms to solve MIP.

In this paper, we analyze the practicability of using exact algorithms to solve this problem. Indeed, existing approaches systematically rely on problem relaxations, decompositions or heuristics. We propose a mathematical model that is solved to near-optimality for instances of reasonable size involving over 70 station yards, over 200 locomotives and a 2-week planning horizon. This scale compares well with the size of a small to medium sized railroad. We next improve the specification by including several valid inequalities, and are then able to solve larger instances with over 1800 locomotives and 600 station yards. Finally, we introduce in the model an uncertainty feature (in the sense of Eggenberg et al., 2011) which allows us to generate solutions that are robust to uncertainty in the fuel consumption parameter. A sensitivity analysis shows that a solution with a small amount of reserve fuel can be implemented for an increase of the cost of about 0.01% when the fuel consumption of the locomotive rises by as much as +10%.

2. Problem Features and Assumptions

In this section, we describe the inputs to the problem and the underlying assumptions considered in this paper. The most basic input is the train schedule that provides the list of trains and time table. The sequence of yards through which a train halts or passes through is also known from the train schedule. Some trains operate daily while some others operate on fewer days of the week. The sequence of yards that a train stops or passes by is usually identical for any train over the different days of its operation. This is not restrictive as the same train with the same origin-destination (OD) can be provided with a different train number (indicator) if the train does not stop or pass by the same sequence of yards. Onward and return journeys are indicated with different train numbers. Each train may run for one or more days. However it is assumed that the train schedule repeats itself over a fixed planning horizon, which for the given data is two weeks.

In addition to train schedule, locomotive assignment plan for trains is also available as an input. This plan assigns locomotives to trains and makes sure that locomotive assignment is feasible and cost effective. The plan generally repeats over the planning horizon, two weeks in the case of our data. So every train-locomotive assignment for the entire planning horizon is considered at the outset. Typically every train may run with one or more locomotives and there would be multiple types of locomotives available to be scheduled for a train, depending on tonnage and horse power requirements. However, we assume that all locomotives of the same type are used and that every train is powered by exactly one locomotive. This assumption is again not restrictive for planning mode, even though it is restrictive for operations mode. It is so because the problem solved by us can be considered as the decomposed version of the problem involving multiple locomotive types or numbers of locomotive at planning stage.

The locomotive refueling problem considered in this paper is motivated by the INFORMS Problem Solving Competition 2010. This problem considers a rail network, a locomotive plan that describes the assignment of locomotives to trains (on particular days) and the train time table. Fueling costs have three distinct components. One relates to actual cost of fuel which is simply calculated by the cost of fuel at the yard multiplied by the amount of fuel filled in the locomotive. Second is a fixed nominal cost for every stop that the locomotive makes for fueling. This cost is independent of the characteristics of the network, yard, locomotive or its schedule. This cost is incurred at every instance when the locomotive halts for refueling and is always the same. Third cost component relates to the cost of holding a fueling truck at a particular yard. We assume that the entire exercise is a precursor to negotiating contracts with vendor companies to hold dedicated fueling trucks at certain yards, every day, all round the year. The costs for locating the fueling truck at a yard or for a locomotive to make a fueling halt need not be same. Note that a fueling halt at a train station yard that has no fueling truck would not only be undesirable but also trite. The objective

of this paper is to determine the refueling halts and placement of fueling trucks across the entire network such that the total costs are minimized. This problem involves a clear trade-off between the purchasing costs of fuel vis-à-vis fueling costs of locomotives at specific yards. On top of this trade-off, the problem of locating fueling trucks is similar to a set covering problem involving the identification of minimal set of yards (to locate these trucks), over the optimal trade-offs between locomotive halting and fuel purchasing costs, that covers all the locomotive paths in the planning horizon.

The assumptions associated with the problem considered here are listed below.

- Amount of fuel consumed by all locomotives is assumed to be known in advance over the entire network. It effectively means that the amount of fuel consumed by the locomotive to run between station yards A and B is deterministic and consistent. It does not change based on weather or wind conditions.
- Since the problem considered here is in planning mode, the assigned locomotives are assumed to be available for all trains on all days throughout the network.
- Train time table is assumed to be followed strictly and scheduling decision does not make any provision for delays or deviations.
- Capacity of the locomotive fuel tank beyond the minimum safety level is known in advance.
- Fueling trucks also have a known uniform capacity limit for the amount of fuel that can be dispensed on a given day. Fueling trucks would normally start the day at full capacity. Of course, the mathematical model presented in the next section can be easily adapted if there are multiple possibilities for the capacity of fueling trucks with different costs.
- Any fueling stop that is neither origin nor destination is an intermediate stop. It is also desired that the number of intermediate fueling stops for a locomotive on a train route is bounded by a given number.

3. Mathematical Model

We would first break the problem of trains and locomotives into a path for locomotives only, discarding the train. Thus the sequence of yards visited by a locomotive on different days would be derived. Let a locomotive visit a sequence of yards across all train routes and all days in the entire planning horizon and be represented by index s such that $s = \{1, 2, 3 \dots S\}$. The sequence of yards s is drawn in such a way that destination yard for every train route is skipped because it is same as the origin yard for the next train route of the same locomotive and also because refueling is not an option at the destination yard. Let j be the index used to identify a locomotive. Let r represent the index for train route (different for every time period, say, day), i be the yard and t be the day number in the planning horizon, varying between 1 and T . This section is divided into a brief description of the decision variables, followed by parameter descriptions and then the objective function definition and finally the constraints.

Decision Variables:

- x_{js} : Flag to represent refueling of locomotive j at the yard appearing in sequence s on its route; Binary
 y_{js} : Amount of fuel in locomotive j at the time of **entering** the yard appearing in sequence s ; Linear
 w_{js} : Amount of fuel filled in locomotive j at the yard appearing in sequence s ; Linear
 z_i : Number of refueling trucks at yard i ; Integer

Known Parameters:

- Param_refuel_{jis}: 1 if locomotive j visits yard i on sequence s and 0 otherwise
Day_{jst}: 1 if locomotive j visits yard sequence s on day t and 0 otherwise
Train_{rjs}: Flag for intermediate yards; 1 if yard sequence s for locomotive j on train route r is Intermediate, 0 otherwise (*pre-processed*)
 d_{js} : Distance between yards appearing in sequence s and next (in case it is the last non-destination yard, then the first yard of the sequence) for locomotive j (*pre-processed*)
rate: Amount of fuel consumed to run one mile
Min_fuel_{js}: Minimum fuel required to reach next yard (= $d_{js} * \text{rate}$)
 c_i : Cost of fuel at yard i
 c_{FIXED} : Fixed cost for refueling
 $c_{CONTRACT}$: Contracting cost of refueling truck over the planning horizon
CAP: Refueling truck capacity
TANK: Locomotive tank capacity
NFP: Maximum number of intermediate fueling yards on a train route

Objective Function:

Objective function has the three cost components. The first term represents that fixed halting cost at any station yard for refueling, second term represents the cost of purchasing fuel at that yard and the third term indicates that cost of contracting z_i number of fueling trucks at yard i .

Minimize Cost

$$\sum_j \sum_s c_{FIXED} x_{js} + \sum_i \sum_j \sum_s Param_refuel_{jis} c_i w_{js} + \sum_i c_{CONTRACT} z_i$$

Constraints:

1. A locomotive j on yard sequence s is refueled if and only if there is a halt at that yard.

$$w_{js} \leq TANK \cdot x_{js} \quad \forall j, s \quad \dots (1)$$

Constraint (1) ensures that a locomotive cannot be refueled if a yard sequence is not a refueling halt.

Given that c_{FIXED} is a non-negative real number, cost minimization objective ensures that the variable

x does not take a value 1 if corresponding variable w is zero, i.e. there would be no unnecessary halt at a station yard, except for the purpose of refueling.

2. Fuel in the locomotive at any time cannot exceed the tank capacity

$$y_{js} + w_{js} \leq TANK \quad \forall j, s \quad \dots (2)$$

3. Fuel conservation in the locomotive before and after crossing a yard sequence s

$$y_{js} + w_{js} - Min_fuel_{js} = y_{js+1} \quad \forall j, s: s \cap \{S\} \quad \dots (3a)$$

$$y_{js:s=\{S\}} + w_{js:s=\{S\}} - Min_fuel_{js:s=\{S\}} = y_{js:s=\{1\}} \quad \forall j \quad \dots (3b)$$

This ensures that the amount of fuel in the locomotive at the time of entering a station yard is the sum of fuel filled at the previous yard and the amount of fuel in the locomotive at that yard minus the fuel burnt to arrive at the present station yard.

4. A locomotive can be refueled in at most NFP intermediate yards along a route (excluding the origin and the destination)

$$\sum_s Train_{rjs} x_{js} \leq NFP \quad \forall j, r \quad \dots (4)$$

5. There is a limit on the amount of fuel that can be filled at each yard every day which is restricted by the capacity of the fueling truck at that yard

$$\sum_j \sum_s Day_{jst} Param_refuel_{jis} w_{js} \leq CAP. z_i \quad \forall i, t \quad \dots (5)$$

6. Bounds on the variables

$$w_{js} \geq 0 \quad \forall j, s \quad \dots (6a)$$

$$y_{js} \geq 0 \quad \forall j, s \quad \dots (6b)$$

$$x_{js} \in \{0,1\} \quad \forall j, s \quad \dots (6c)$$

$$z_i \in Integer, \geq 0 \quad \forall i \quad \dots (6d)$$

The above MIP formulation is theoretically adequate to find an optimal solution. However when the problem instance scale is large, the run times are expected to be higher on a standard solver. Details of the run time and results are indicated in the next section. One of the factors that makes the problem inherently complex is that the amount of starting fuel for each locomotive actually has several possibilities in the same optimal solution. To illustrate with an example, let us say a locomotive visits six yards sequentially during the planning horizon and the amount of fuel needed to go between each is 800 gallon. If the locomotive capacity is 4500 gallons and fifth yard is cheapest for fueling, the starting fuel can be any amount between 4000 and 4500 gallons in the optimal solution. Nourbakhsh and Ouyang (2010)

tackled this problem by ensuring that only a finite number of alternatives are considered in the optimal solution through a proposition that is proved logically.

We cannot use the same proposition even though the problem characteristics are similar, due to the constraint pertaining to maximum fuel that can be dispensed at a yard for a specific time period and a restriction on the maximum number of refueling points for every train. According to proposition 3 of Nourbakhsh and Ouyang (2010):

“If a locomotive purchases fuel at two fixed fueling stations s_1 and s_2 (not necessarily adjacent along the route) but no emergency fuel in between, then there exists an optimal solution in which the locomotive either departs s_1 with a full tank, or arrives at s_2 with an empty tank.”

Note that the problem considered by us has a volume restriction on the amount of fuel that can be dispensed at each yard. Additionally there are also restrictions on the number of intermediate (non-origin, non-destination) refueling station yards. So there could be an optimal solution where the locomotive could not leave s_1 with a full tank (due to limited availability of fuel at the yard), but arrives at s_2 with a non-empty tank (even as it is not profitable to refuel at s_2 because of higher fuel price). A counter example is shown in the appendix to prove that the proposition does not hold for this problem.

4. Tightening the MIP Formulation by Introducing Cuts

In this section, we introduce some valid inequalities to improve the model performance and efficiency.

If a station yard does not have a contracted refueling truck, it is not possible to fuel the locomotive at that yard. It would effectively mean that the x variable for a particular locomotive and stop sequence has to be less than or equal to the z variable corresponding to the yard in the optimal solution.

$$\sum_j \sum_s Param_refuel_{jis} x_{js} \leq z_i \quad \forall i \quad \dots (7)$$

This inequality is indeed valid and would also reduce the feasible space.

Given a fixed capacity for the locomotive fuel tank, it is obvious that the tank gets refueled in one of the subsequent finite number of yards before it runs out of fuel to even reach the next yard. Let us consider the following example where a locomotive travels through 11 station yards sequentially. Refer to Figure 1 for illustration.

Let us assume that the locomotive tank capacity is 4500 gallons while it takes 600 gallons to travel between stop sequence 1 and 2, 700 gallons to travel between stop sequence 2 and 3, and so on. If the locomotive leaves stop sequence 1 with a full tank, it can travel from stop sequences 2 to 7 without

refueling. But it cannot reach stop sequence 8 without a refuel at one of the intermediate yards. Thus it is imperative that the locomotive is (re)fueled in **at least** one stop sequence among 2 to 7. Next we consider stop sequence 2. If we assume that the locomotive leaves stop sequence 2 with full tank, it can travel between stop sequences 3 to 8, but cannot continue to the next stop sequence 9 without refueling. The same relationship can be deduced for every locomotive and every stop sequence.

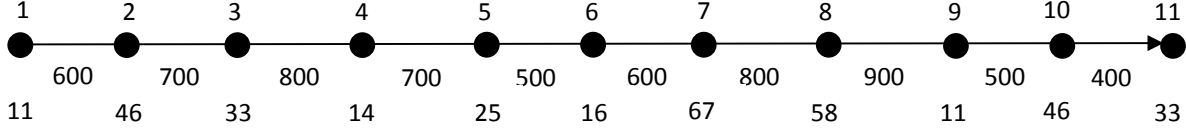


Fig 1: Synoptic map of a locomotive travelling from stop sequence 1 through 11.

Amount of fuel required to reach the next stop and the actual yard numbers are shown below

For every yard sequence $s_k \in s$, there exists a finite set of station yards $\mathfrak{S}_k \in \{s_k + 1, \dots, s_k + n\}$ such that the locomotive must be refueled in at least one of them to be able to continue the journey. The following cut represents the introduction of this constraint.

$$x_{j,s_k+1} + \dots + x_{j,s_k+n} \geq 1 \quad \forall j, s_k \in s: \{s \cup (S+1 \equiv 1 \in s)\} \quad \dots (8)$$

Let us consider the same example as above. In Figure 1, we now consider the physical yard numbers instead of stop sequences. Note that the condition applicable for stop sequences is also applicable for the yard numbers. Thus a locomotive leaving yard 11 **must** get (re)fueled in **at least one** of the yards – 46, 33, 14, 25, 16, or 67.

Thus, without loss of generality, it can be assumed that these yard sequences represent the set of physical yards, $I_{\mathfrak{S}_k} \in \{i_p, \dots, i_q\}$. Note that the cardinality of the set $I_{\mathfrak{S}_k}$ would be lesser than or equal to the cardinality of the set \mathfrak{S}_k as some yard sequences may represent the same yard.

$$z_{i_p} + \dots + z_{i_q} \geq 1 \quad \forall i_p, \dots, i_q \in i \quad \dots (9)$$

As already mentioned earlier, the objective function has three distinct cost components for the cost of fuel, cost of halting at a station yard and the cost of contracting a fuel truck at that yard. The objective is a strictly non-decreasing function. It is so because the variables themselves are non-negative and the cost coefficients corresponding to the cost of fuel, cost of halting and the cost of the fuel truck cannot be negative either. While the cost of fuel purchase and the cost of halting at a station is locomotive specific, the cost of contracting the fuel truck is dependent on the entire network. For example, if there was no cost for contracting a fueling truck or that unlimited fuel was always available at every station yard, the problem of minimizing fuel purchase cost and halting cost can be decomposed at locomotive level. Thus the optimal cost of minimizing fuel purchase decision and the cost of halting for each locomotive is same as the optimal cost obtained by considering all these locomotives together.

Decomposed problem for each locomotive, being a small instance, is usually solved to optimality very quickly (within a few milliseconds, usually) by a commercial solver. Optimal costs for purchasing fuel and halting obtained for the decomposed problem can be used as a very tight lower bound of the corresponding cost components for each locomotive in the original problem.

Thus the first two terms of the objective function are minimized for each locomotive separately and MIP is solved for constraints (1) to (4) and (6a) to (6d). Let the optimal locomotive halting cost parameter for every locomotive at every yard sequence be x_{js}^* and the optimal amount of fuel filled at yard sequence s for locomotive j be w_{js}^* . Thus for every locomotive, the total fuel purchasing cost in the entire planning horizon will be $\sum_i \sum_s Param_refuel_{jis} c_i w_{js}^*$ and the total cost of fuel halting would be $\sum_s c_{FIXED} x_{js}^*$. Therefore, we can add the following two constraints that give a lower bound on locomotive-level fuel purchasing cost and halting cost for the original problem.

$$\sum_i \sum_s Param_refuel_{jis} c_i w_{js} \geq \sum_i \sum_s Param_refuel_{jis} c_i w_{js}^* \quad \forall j \quad \dots (10)$$

$$\sum_s c_{FIXED} x_{js} \geq \sum_s c_{FIXED} x_{js}^* \quad \forall j \quad \dots (11)$$

The above constraints would ensure a better lower bound and possibly yield quicker solution. It is also possible to get a lower bound on the total cost of truck contracting and fuel purchasing by ignoring the x -variable in the objective function.

5. Problem Instances and Results

We are given a real-life scaled down version of the actual problem of optimal locomotive refueling at a major US railroad through the INFORMS RAS 2010 Open Competition. We would apply the model and solution technique described in sections 3 and 4 above to illustrate the actual performance on the case study problem. This problem has 214 locomotives operating 214 different trains. Some trains are daily and some others are less frequent. However a timetable, which repeats on a weekly basis and the assignment of locomotives to trains, is provided. Even though the time table repeats over a week, locomotive-assignment schedule has a two-week window. Thus we choose a planning horizon of two weeks for our problem as well. Cost of fuel ranges from \$2.90 to \$3.56 per gallon across different station yards. Locomotive tank capacity is 4500 gallons. Cost of halting for refueling is fixed at \$250 per halt and the cost of contracting a fueling truck with a capacity of 25,000 gallons per day is \$4000 per week.

However we are allowed to provide multiple fueling trucks at the same station yard if it makes price sense.

The problem pre-processing is done using MS-Excel and C-programming. ILOG-Cplex is used as the commercial solver to solve the MIP. The given problem instance is run on a 64-bit Linux server with 3.0 GHz processor speed and 4 GB RAM. The formulation without the addition of MIP cuts and tightening bounds produces a solution of about \$11.41 mil with an optimality gap of over 1.5% in about 10 min. The run time to generate the lower bounds x_{js}^* and w_{js}^* for constraints (10) and (11) of the model is small and is usually within 1 minute for all the locomotives. The final model with all cuts and improved bounds produces a solution of \$11,399,670.88 with an optimality gap of 0.08% in less than 10 mins. The performance of the model with the introduction of the various MIP cuts is shown below. All models were run for a cut-off time of 10 mins. The results are shown in table 1.

Model Name	Constraints Included	Time Limit (s)	Solution (mil \$)	Optimality Gap (%)
Base	(1) – (6)	600	11.41150	1.53%
MIPCut#1	(1) – (7)	600	11.40419	0.85%
MIPCut#2	(1) – (9)	600	11.40068	0.21%
MIPCut#3	(1) – (11)	600	11.39967	0.08%

Table 1: Model Performance

We ran the model with all MIP cuts for about 15 hours. The model could not prove the optimality but we were left with a small absolute optimality gap of less than \$100.

We are now focusing on the performance of our model on larger problem instances. Due to lack of real large problem instance, we used the existing network framework. We created network instance of two times, four times and eight times the size of the initial instance by increasing the number of nodes and arcs, thereby creating mirror networks. Our data relating to arc distances was kept the same even in the new “expanded” network. We also proportionally expanded the other features of the problem – such as the station yards, trains and locomotives. All the costs – cost of fuel, cost of halting, cost of contracting fueling truck – as well as the capacity of fueling truck was kept the same. As we know the properties of the new “expanded” network, we also know that the solutions for the bigger sized networks have to be proportionally larger. Thus we know that the best solution that can be expected from the solver for a double sized network would be \$ 22799341.76. The results of the models of larger sized instances are given in table 2 below. The last column indicates the gap with respect to the best known solution extrapolated for the larger network.

Network Size	Model Name	Constraints Included	Time Limit (s)	Solution (mil \$)	Optimality Gap (%)	Best Known Solution Gap (%)
Double	Base	(1) – (6)	600	22.84456	1.89%	0.20%
Double	MIPCut#1	(1) – (7)	600	22.82189	1.06%	0.10%
Double	MIPCut#2	(1) – (9)	600	22.81712	0.48%	0.08%
Double	MIPCut#3	(1) – (11)	600	22.80818	0.36%	0.04%
Quadruple	Base	(1) – (6)	600	45.74321	2.12%	0.32%
Quadruple	MIPCut#1	(1) – (7)	600	45.69412	1.34%	0.21%
Quadruple	MIPCut#2	(1) – (9)	600	45.68753	0.66%	0.19%
Quadruple	MIPCut#3	(1) – (11)	600	45.67923	0.47%	0.18%
Eight Times	Base	(2) – (6)	600	91.72433	2.56%	0.58%
Eight Times	MIPCut#1	(2) – (7)	600	91.65912	1.73%	0.51%
Eight Times	MIPCut#2	(2) – (9)	600	91.63665	1.01%	0.48%
Eight Times	MIPCut#3	(2) – (11)	600	91.61742	0.79%	0.46%

Table 2: Model Performance on Larger Problem Instances

6. Dealing with Uncertainty

In our optimization strategy, we have primarily aimed to minimize costs while making sure that the generated solution is theoretically “feasible”. One of the strongest assumptions in this paper relates to assumption of deterministic fuel consumption throughout the locomotive network. However it is well known that this assumption does not make any sense in practice. In real-life, locomotives must be having additional capacity to carry reserve fuel throughout their journey. However any discussion on the amount of reserve fuel has been left out in the discussion so far. We would consider the implication of this assumption and suggest a method to overcome the same in this section.

Eggenberg et al. (2011) have provided an interesting framework to enhance the robustness of the solution in the face of data uncertainty without increasing the model complexity. We would adapt a similar philosophy to our context. While the amount of fuel required to travel between two given yards is primarily dependent on the distance between the two yards, which is known in advance, there are a number of other minor factors which might cause some variance. Fuel requirement for a locomotive varies to travel the same two yards might vary due to changes in train speeds, braking situations due to signals, atmospheric pressure, wind conditions and temperature. Curiously, when we analyzed the optimal solution for the base network, we found that if the locomotive consumed even 1% more fuel at a particular section (between two yards), there might have been 755 occasions (more than two-thirds of refueling

halts) when the train would have run out of fuel before arriving at the next yard with a fueling truck. While this measure is indicative of the extent to which the theoretical model has been optimized, it also indicates that this solution could be of little interest to the practitioners.

To make the optimal solution practically implementable, we suggest **maximizing** the minimum fuel in the locomotive tank at any point in time which can be utilized when the locomotive draws more than normal fuel in a particular route. The key here is to increase the minimum fuel in the locomotive tank **without adversely impacting the cost**. Let us say that such minimum fuel for locomotive j be y_min_j . The objective function for the model is rewritten as:

Minimize Cost

$$\sum_j \sum_s c_{FIXED} x_{js} + \sum_i \sum_j \sum_s Param_refuel_{jis} c_i w_{js} + \sum_i c_{CONTRACT} z_i - \sum_j \alpha \cdot y_min_j$$

where α is a weight (importance) factor for minimum fuel in the objective function which must maximize the fuel without increasing the cost, which is represented by the first three terms. As a result, it is necessary to choose a very small value of α that would increase y -variable without affecting the cost. For our case-study, we chose α as 0.001 which ensures that the cost does not change while y_min_j is maximized.

In addition to constraints (1) – (11), the following additional constraint would ensure that y_min_j is indeed maximized.

$$y_min_j \leq y_{js} \quad \forall j, s \quad \dots (12)$$

The optimizer is again run on the modified model with a cut-off time of 10 min. While the optimality gap now increases to 0.12%, the optimal costs still remain \$11399670.88. But significantly, this change in the objective function results in a new solution that reduces the number of stock-out occasions when the locomotive consumes more fuel on every section to 154 (from the previous 755). This method makes it possible to obtain a robust solution by changing the refuel amount and fueling locations without increasing the cost.

While this solution is more robust than the solution obtained in the previous section, it is still not practical to be implemented for the real-life instance because of the high chance of stock-out even if there is a small variation in the fuel consumption of the locomotive. To make the solution less sensitive to the fuel consumption of the locomotive, it is necessary to carry some additional fuel as reserve. Forcing locomotives to carry some reserve fuel would escalate the costs, but the solution would be practical as the chance of in-transit fuel stock-outs can be nullified (or minimized). Let us say that the amount of fuel variation (up-side) for each locomotive can be up to β . The value of β can be determined by past

experience and can be chosen differently for each locomotive and section. The mathematical model can be re-written to accommodate this variation by adding the following constraint to the model represented by equations (1) to (12) along with the new objective that outputs a robust solution.

$$y_{js} + w_{js} - (1 + \beta) \cdot \text{Min_fuel}_{js} \geq 0 \quad \forall j, s \quad \dots (13)$$

In this model, an expectation that the fuel consumption can be up to 10% more beyond the normal would mean that β would take a value 0.1. The more we increase the value of β , the more we tend to increase the reserve fuel in locomotives and hence costs. We find that if the fuel consumption increases by up to 10% on any section in the network, the corresponding increase in cost of maintaining this additional reserve would be \$1178.62 – which is an increase of about 0.01% of the total costs. Thus, interestingly the cost of providing a robust solution is not too high for this problem instance. The relationship between β and the increased costs when the fuel consumption of the locomotives vary from +1% to +30% is shown in the pareto curve in figure 2.

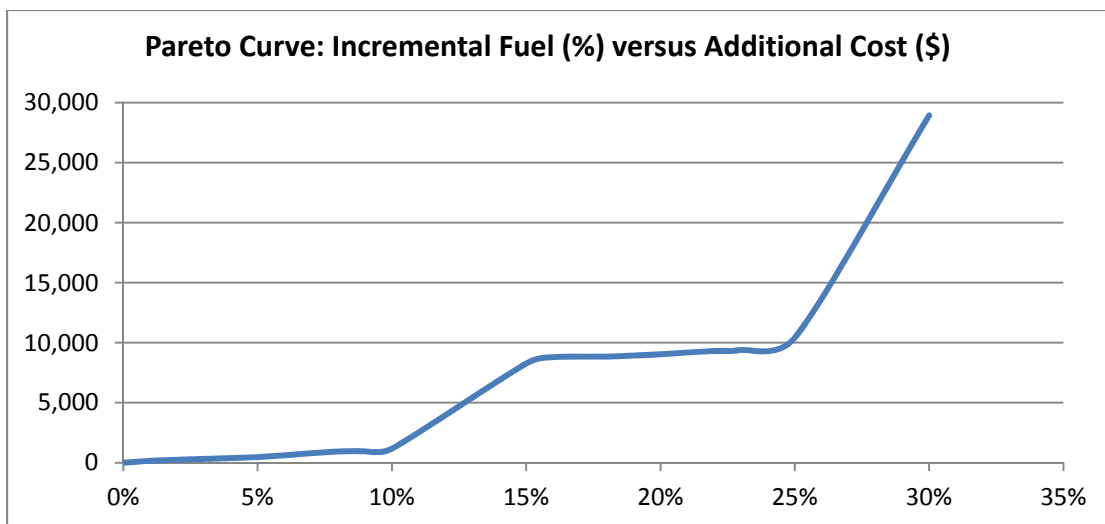


Fig 2: Additional cost incurred to optimize the incremental fuel consumption

7. Future Research Directions

In the conclusion, we have shown that realistic instances of the problem can be solved to near-optimality with exact methods, thanks to adequate valid inequalities. We have proved by this work that railroad companies must look beyond heuristics and crude solution methods as the exact approaches are shown to work well for even much larger network instances. In our work, we have also shown that

variances in fuel consumption by the locomotives can be handled at minimal incremental costs by maintaining a certain reserve fuel.

While the concept of adding valid inequalities as an operations research technique has been studied for discrete optimization problems extensively in the theoretical context, our application shows its significance in practice. In addition to the railroad industry, this research would clearly have relevance in the context of airlines as well. Aircrafts fly on predetermined routes and they do have long term contract with fuel vendors at certain airports. The price of the jet fuel also varies across airports. Cost of fueling aircraft fleet can be optimized by building an appropriate MIP model and using MIP cuts for faster solution convergence. Apart from the applications in fueling, the model and the concept can also be extended to other scenarios that have fixed and operating costs in the problem structure. Cost of ordering goods from multiple vendors selling the same commodity with different inventory levels and warehouse locations and varying transportation costs can be modeled and solved as our problem.

The next steps in the research on this problem should focus on incorporating robustness of fueling decisions. We have suggested one method of capturing robustness by maximizing the minimum reserve fuel that the locomotive must always carry to avoid stock-outs during its journey. It would be interesting to study other robust optimization techniques and report their performance for this problem.

Another point to be noted is that the term “robustness” has much wider connotations. While we are specifically discussing about the fueling decision robustness due to variances in fuel consumption, we have not touched upon other considerations which might also impact “tight and close to optimal” fueling plans. One such scenario can arise when the fueling truck fails to reach a particular yard on a given day, thereby creating a possible crises situation. It may be worthwhile to explore the option of purchasing emergency fuels in such situations and accounting for the same in the costs during the modeling stage.

Acknowledgements

We would like to thank Kamlesh Somani and Juan Morales for conducting the INFORMS Railway Application Section competition 2010 and providing us with relevant practical data for the problem.

References

[1] Eggenberg, N., Salani, M., and Bierlaire, M. (2011). Uncertainty Feature Optimization: an implicit paradigm for problems with noisy data, *Networks*57(3):270-284. doi:10.1002/net.20428

- [2] Khuller, S., Malekian, A., Julian, M., 2007. To fill or not to fill: the gas station problem. In: Arge, L., Hoffmann, M., Welzl, E. (Eds.), *Lecture Notes in Computer Science*, vol. 4698. Springer-Verlag, Berlin Heidelberg, pp. 534–545.
- [3] Kuby, M., Lim, S., 2005. The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences* 39 (2), 125–145.
- [4] Kuby, M., Lim, S., 2007. Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. *Networks and Spatial Economics* 7 (2), 129–152.
- [5] Kuby, M., Lim, S., Upchurch, C., 2005. Dispersion of nodes added to a network. *Geographical Analysis* 37 (4), 384–409.
- [6] Lim, S., Kuby, M., 2010. Heuristic algorithms for siting alternative-fuel stations using the flow-refueling location model. *European Journal of Operational Research* 204 (1), 51–61.
- [7] Nag, B., Murthy, K. G., 2010. Locomotive Fueling Problem (LFP) in Railroad Operations, Technical Report, University of Michigan (available at <http://www-personal.umich.edu/~murty/>)
- [8] Nourbakhsh, S.M., Ouyang, Y., 2010. Optimal fueling strategies for locomotive fleets in railroad networks. *Transportation Research Part B* 44, 1104–1114.
- [9] Problem Solving Competition 2010, Railway Applications Section (RAS) of INFORMS, (available at <http://www.informs-ras.org/problem.htm>)
- [10] Stroup, J.S., Wollmer, R.D., 1991. A fuel management model for the airline industry. *Operation Research* 40 (2), 229–237.
- [11] Upchurch, C., Kuby, M., Lim, S., 2009. A capacitated model for location of alternative-fuel stations. *Geographical Analysis* 41 (1), 127–148.
- [12] Wang, Y.W., Lin, C.C., 2009. Locating road–vehicle refueling stations. *Transportation Research Part E* 45 (5), 821–829.
- [13] Zouein, P.P., Abillama, W.R., Tohme, E., 2002. A multiple period capacitated inventory model for airline fuel management: a case study. *Journal of the Operational Research Society* 53 (4), 379–386.

Appendix

Consider a locomotive assigned to a train with three stations, starting at the origin station and passing by two intermediate stations and returning to the same origin station at the end of the planning horizon. Fuel requirement between origin and the first intermediate station is 1500 gallons, first intermediate station to second is 2000 gallons, and second intermediate station to destination is 1500 gallons. The prices of fuel at the three stations are \$3, \$2 and \$3 respectively. Let the locomotive tank capacity be 4500 gallons and the locomotive can be refueled at only one intermediate station. While fuel availability at origin and second intermediate station are 10,000 gallons, it is only 4000 gallons at the first intermediate station. It could be possible that the actual fuel availability is much higher, but other locomotives passing through

the yard also require a share of the cheap fuel dispensed at this yard, thereby reducing the availability for this locomotive. Table below shows the possible optimal fueling sequence and starting fuel range for this locomotive.

Yard Sequence	Yard Number	Yard type	Min fuel to reach next yard	Cost of fuel	Starting fuel range	Fuel filled	Exit fuel range	Max available fuel at the yard
1	Y1	Origin	1500	3	500 - 1000	1000	1500 - 2000	10000
2	Y2	Intermediate	2000	2	0 - 500	4000	4000 - 4500	4000
3	Y3	Intermediate	1500	3	2000 - 2500	0	2000 - 2500	10000
1	Y1	Origin						

Table 1: Range of starting fuel at different yards for optimal solution

It can be noted that when the arriving fuel at Y2 is 0, then locomotive tank is filled up to 4000 (not full tank) and reaches Y1 to refuel when the fuel is still non-zero (500 gallons). Similarly, when the arriving fuel at Y2 is 500 gallons, locomotive can leave Y2 with full tank, arrive Y1 with 1000 gallons fuel, leave Y1 with 2000 gallons fuel (not full tank) and refuels at Y2 when the fuel level in non-zero (500 gallons). Therefore, there is at least one cycle in the example above where proposition 3 of Nourbakhsh and Ouyang (2010) does not hold.