

Optimizing the aerodynamic efficiency of intermodal freight trains

Yung-Cheng Lai^{a,*}, Christopher P.L. Barkan^a, Hayri Önal^b

^a *Railroad Engineering Program, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 N. Mathews Avenue, B118 Newmark Lab., Urbana, IL 61801, USA*

^b *Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 1301 W Gregory Dr., 306 Mumford Hall, Urbana, IL 61801, USA*

Received 8 August 2006; received in revised form 24 April 2007; accepted 17 May 2007

Abstract

We develop an aerodynamic loading assignment model for intermodal freight trains based on an integer-programming framework to help terminal managers make up more fuel-efficient trains. This is the first use of optimization modeling to address the aerodynamics and energy efficiency of railroad intermodal trains. Several recommendations regarding railway equipment use, operations, and policy are proposed to improve fuel-efficiency and reduce emissions from intermodal transportation. Analysis of one major railroad intermodal route reveals the potential to reduce fuel consumption by 15 million gallons per year with corresponding savings of \$28,000,000. Greater benefits are possible through broader implementation of the model.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Energy efficiency; Aerodynamics; Fuel use; Railroad intermodal transportation; Loading assignment; Urban freight

1. Introduction

Intermodal (IM) freight is the largest segment of the US railroad freight transportation business, and is the fastest growing portion (nearly 80% in the past 15 years) (Gallamore, 1998; Association of American Railroads (AAR), 2005). However, IM trains are generally the least fuel efficient trains. This inefficiency is due to the physical constraints imposed by the combination of loads and the railcar design (Engdahl et al., 1987a,b; Lai and Barkan, 2005). In the pair of photographs shown in Fig. 1 contrast the close spacing of the hopper cars with the much larger gaps due to the empty slots in the IM train. The aerodynamic drag coefficient of a typical IM train can be as much as 25% higher than that of the fully loaded coal train, and this difference in resistance increases exponentially with speed (American Railway Engineering and Maintenance-of-Way Association, 2001; Bertin, 2001; Hay, 1982). IM trains commonly operate at speed of

* Corresponding author. Tel.: +1 217 721 4525; fax: +1 217 333 1924.
E-mail address: lai3@uiuc.edu (Y.-C. Lai).



Fig. 1. Pictures of: (a) general coal cars; (b) trailers on spine cars.

70 mph making them the fastest freight trains in North America, thus aerodynamic drag is a particularly important factor affecting their fuel consumption.

Lai and Barkan (2005) conducted a series of analyses to compare both the relative and absolute effects of different loading patterns and operating practices on train make-up and energy efficiency. They found that aerodynamic characteristics significantly affect IM train fuel efficiency; and, a train can be more efficiently operated if loads are assigned not only based on slot utilization but also by properly assigning loads to cars, which they referred to as “slot efficiency”. Depending on the particular train configuration, train resistance for a fully loaded train can be reduced by as much as 27%, and fuel savings by 1 gallon per mile per train, simply by better matching loads and railcars.

Lai and Barkan’s previous work (2005) quantified the aerodynamic and energy penalties of specific load and car combinations under idealized conditions. They did not consider the actual make-up of train consists or the wide variety in available loads and car types that a terminal manager must contend with in trying to implement more energy efficient loading practices. Lai et al. (2007) describe a wayside machine vision system that automatically monitors the gap lengths between IM loads on passing trains so the railroad can evaluate the aerodynamic efficiency of their loading pattern. However, no previous work has addressed the question of how to select among the wide variety of loads and railcars actually available to load aerodynamically efficient trains. This is an essential element of achieving the potential fuel and costs savings. In this paper, we develop an aerodynamic loading assignment model (ALAM) using an integer programming (IP) framework to optimize aerodynamic efficiency under various constraints regarding loading assignments. The model can help terminal managers load trains more efficiently and can be incorporated into software used to automate or expedite the loading process inside IM terminals.

Previous researchers have considered various other aspects regarding optimization of the loading process and equipment utilization. Feo and Gonzales-Velrade (1995) proposed an integer-linear programming model to maximize the utilization of trailers to railcar hitches. Powell and Carvalho (1998) developed a dynamic model to optimize the flow of flat cars over a network. Corry and Kozan (2006) presented an assignment model to dynamically assign containers to IM trains so as to minimize excess handling time and optimize the weight distribution of the train. Each of the above studies focused on certain types of IM loads or railcars. However, none of them considered the energy efficiency of IM train loading. In this paper, we present the first application of optimization techniques to improve the energy efficiency of IM trains. The proposed model can deal with all types of IM loads (11 different types of trailers and containers), and railcars (hundreds of different types of well, spine, flat cars) operated in North America (TTX Company, 1999).

This study is particularly timely in light of increasing fuel prices and their impact on industry operating costs, as well as the need to conserve energy and reduce greenhouse gas emissions. In 2005, the major North American railroads spent over \$8 billion on fuel in the United States making it their second largest operating expense. From 2002 to 2005, North American railroad fuel cost doubled, and since 1999 it is up by nearly a factor of three (Association of American Railroads, 2006). This trend is impacting railroads all over the world, consequently fuel efficiency is more important than ever (Stodolski, 2002; BNSF Railway Company, 2004a; International Union of Railways (IUC), 2004). Investigation of options to improve aerodynamic efficiency is a promising avenue with widespread potential economic and environmental benefits for rail freight transportation efficiency (Association of American Railroads (AAR), 1981 and Association of American Railroads (AAR), 1987; Smith, 1987).

2. Intermodal rail terminal operations

At IM terminals, managers assign containers and trailers of a variety of lengths to available well, spine or flat cars (BNSF Railway Company, 2004b; Union Pacific Railroad Corporation, 2004). Railroad IM business in North America is different from the general carload freight business; IM business often competes directly with trucks and as a consequence is very time sensitive. Because of this, railroads try to avoid intermediate switching and stops on most IM trains. In this study, we used the IM operation of the BNSF Railway's route between Chicago and Los Angeles (LA) (aka "the Transcon") as the basis for our analyses. This is one of the busiest IM corridors in North America and approximately 80% of the IM trains on this route have no intermediate operations. Of the remaining 20%, most have no more than two intermediate stops and these are generally close to the final destination (Utterback, 2006). Given that for most trains there is little or no container shifting occurring en route, the initial loading pattern will be the principal factor affecting their aerodynamic performance.

IM loads, i.e. trailers or containers, range in length from 20 to 57 ft (Muller, 1999; TTX Company, 1999). There is considerable variety in the design and capacity of IM railcars with different numbers of units and slots, and thus loading capabilities. An IM railcar may have one or more units permanently attached to one another (via articulation or drawbar). A unit is a frame supported by at least two trucks, providing support for one or more platforms (a.k.a. slots). For example, Fig. 2a shows an articulated 3-unit well car, and Fig. 2b is a 5-unit spine car. A platform (or slot) is a specific container/trailer loading location. As a result, each well-car unit has two slots because of their ability to accommodate two containers, one stacked on the other (a.k.a. "double stack"), and each spine-car unit has one slot (Fig. 2).

There are also a number of safety-related loading rules and various feasible and infeasible combinations of IM load and car configurations. Because IM cars in a train are not generally switched in and out at terminals, managers primarily control the assignment of loads but not the configuration of the equipment in a train. Consequently, it is reasonable to treat train make-up as given. Terminal managers often use computer software (Optimization Alternatives Ltd., Inc., 2005) as a decision-making tool to assist them in complying with loading rules; nevertheless load assignment is still a largely manual process in which aerodynamic efficiency is not considered.

3. Methodology

To develop ALAM we need to quantify the IM train aerodynamic characteristics in order to incorporate them into the integer programming model for optimal loading.

3.1. Evaluation of intermodal aerodynamic efficiency

The principal metric currently used to measure the efficiency of IM train loading is "slot utilization" (Burriss, 2003). It measures the percentage of available slots on IM cars that are used for loads. Slot utilization typically does not take into account the size of the space compared to the size of the load. Although perfect slot utilization indicates maximal use of available spaces, it is not intended to, nor does it ensure, that IM cars are loaded to optimize their aerodynamic characteristics and hence maximize energy efficiency. Two trains may have identical slot utilization, but significantly different energy efficiency due to different loading patterns and consequent aerodynamic resistance (Lai and Barkan, 2005).

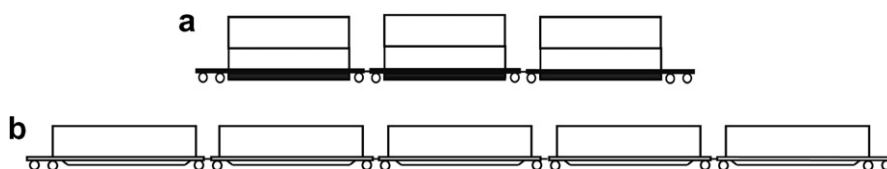


Fig. 2. (a) A 3-unit well car with 6 slots; (b) a 5-unit spine car with 5 slots.

Aerodynamic drag is known to be a major component of train resistance, particularly at high speeds (Hay, 1982; American Railway Engineering and Maintenance-of-Way Association, 2001). The Association of American Railroads (AAR) supported research on wind tunnel testing of rail equipment, including large-scale IM car models (Gielow and Furlong, 1988). The test results were used to develop the Aerodynamic Subroutine of the AAR’s Train Energy Model (TEM) (Drish, 1992). These experiments showed that the gap length between IM loads, position-in-train, and yaw angle of wind are three important factors affecting train aerodynamics (Engdahl, 1987). Although yaw angle of wind is important at specific locations, over a long route this effect will tend to be canceled out by winds in all directions. Consequently, this study focuses on the first two factors, namely gap length and position-in-train effects.

The greater the gaps between loads, the larger the aerodynamic penalty because closely-spaced containers or trailers behave as one long load. In contrast, loads spaced 72" or more apart, behave as distinct objects as boundary layers on their surfaces are reinitialized (Engdahl et al., 1987a). The wind tunnel tests showed that the lead locomotive experiences the highest drag due to headwind impact. After the head end, resistance declines until about the 10th unit in the train, after which drag remains nearly constant for the remaining units. Therefore, minimizing the total gap length and placing loads with shorter gaps near the front of the train will result in lower aerodynamic resistance. Consequently, the objective of optimal loading can be stated as minimization of the “total adjusted gap length (TAGL)” within the train. The adjusted gap length is defined as the gap length weighted by the position-in-train effect, where the weight associated with each unit gets smaller as the unit gets farther from the head end. The relationship between aerodynamic resistance and position-in-train effect is represented in the following equation derived from wind tunnel testing (Engdahl, 1987):

$$C_{DA} = 14.85824e^{-0.29308k} + 9.86549e^{-0.00007k} + 10.66914 \tag{1}$$

where k is the unit position in the train and C_{DA} is the drag area which represents the aerodynamic resistance in ft^2 . The adjusted factor associated with each gap can be computed by dividing the drag area of a given unit by the drag area of the 100th unit (Table 1).

3.2. Aerodynamic loading assignment model (ALAM)

The following notation is used in the algebraic model: $I = \{i\}$ is an index referring to the type or size of the load (for instance, $i = 40C$ represents a 40' container, $i = 48T$ represents a 48' trailer, etc.). C_L and T_L are a subset of I representing containers and trailers, respectively. We group loads of the same type together and denote each load of type i with $j = 1, 2, 3, \dots, J_i$, where J_i is the total number of loads of type i (for instance, $J_{48T} = 10$ means that there are ten 48'-trailers in the storage area). The symbol k ($k = 1, 2, 3, \dots, N$) defines the position of each unit in the train. For instance, $k = 1$ corresponds to the first IM unit of the train, $k = 2$ corresponds to the second unit, etc. The slot position in every unit is denoted by p , where $p = 1$ represents the upper (top) slot (platform) in a well-car unit or the (only) slot in a spine-car or flat-car unit, and $p = 2$ represents the lower (bottom) slot in a well-car unit (see Fig. 3).

Table 1
Adjusted factor of each gap in the train

k	Drag area (ft^2)	Adjusted factor
1 (locomotive)	31.618	1.5449
2	28.801	1.4073
3	26.700	1.3046
4	25.133	1.2280
5	23.963	1.1709
6	23.091	1.1283
7	22.440	1.0964
8	21.954	1.0727
9	21.591	1.0550
10	21.320	1.0418
100	20.466	1.0000

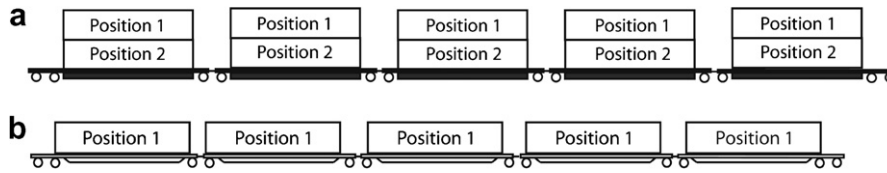


Fig. 3. The available slots in: (a) 5-unit well car; (b) 5-unit spine car.

The following symbols represent the parameters used in the model: A_k is the adjusted factor of the k th gap (see Table 1), where $A_1 > A_2 \dots > A_N$; U_k is the length of the k th unit; δ_k indicates the type of the k th unit, where $\delta_k = 1$ when the unit is a well-car unit, and $\delta_k = 0$ otherwise; L_i is the length of an i th type load; Q_{kp} is the length limit of position p in the k th unit; w_{ij} is the weight of the j th load of type i ; W_k is the weight limit for the k th unit; and R_{ipk} is a three dimensional matrix for loading capabilities of each slot, where $R_{ipk} = 1$ if the i th type of load can be assigned to position p in unit k , otherwise it is 0. Finally, Φ represents an arbitrarily specified large number introduced for modeling purposes as will be explained in the model description below.

Two sets of binary decision variables are included in the IP model. The first variable is denoted by y_{ijpk} where

$$y_{ijpk} = \begin{cases} 1, & \text{if the } j\text{th load of type } i \text{ is assigned to position } p \text{ in the } k\text{th unit} \\ 0, & \text{otherwise} \end{cases}$$

The second binary variable, denoted by x_k , determines whether the position 1 (top slot) in a well unit can be used, namely:

$$x_k = \begin{cases} 1, & \text{if top platform of the } k\text{th unit can be used} \\ 0, & \text{otherwise} \end{cases}$$

According to the loading rules, position 1 of a well-car unit (top slot) can be used when the bottom slot is filled by containers whose total length is at least 40' (Association of American Railroads (AAR), 2004).

The loading problem is formulated here as a linear integer programming model. This model minimizes aerodynamic resistance of an IM train (thus the fuel consumption), which is assumed to be a function of the train's total adjusted gap length, subject to the train characteristics and loading possibilities for a given set of loads. The algebraic expression is given below:

$$\text{Min } 0.5 \times \left\{ A_1 \left(U_1 - \sum_i \sum_j y_{ij11} L_i \right) + \sum_{k=1}^N A_{k+1} \left[\left(U_k - \sum_i \sum_j y_{ij1k} L_i \right) + \left(U_{k+1} - \sum_i \sum_j y_{ij1k+1} L_i \right) \right] \right\} \tag{2}$$

The constraints of the model are as follows:

$$\sum_p \sum_k y_{ijpk} R_{ipk} \leq 1 \quad \forall i, j \tag{3}$$

$$y_{ijpk} \leq R_{ipk} \quad \forall i, j, p, k \tag{4}$$

$$40 - \sum_{i \in C_L} \sum_j y_{ij2k} L_i \leq \Phi(1 - x_k) \quad \forall k(\text{such that } \delta_k = 1) \tag{5}$$

$$\sum_{i \in C_L} \sum_j y_{ij1k} \leq x_k \quad \forall k(\text{such that } \delta_k = 1) \tag{6}$$

$$\sum_{i \in T_L} \sum_j y_{ij2k} \leq 2 \times \left(1 - \sum_{i \in C_L} \sum_j y_{ij1k} \right) \quad \forall k(\text{such that } \delta_k = 1) \tag{7}$$

$$\sum_i \sum_j \sum_p y_{ijpk} w_{ij} \leq W_k \quad \forall k \tag{8}$$

$$\sum_i \sum_j y_{ijpk} L_i \leq Q_{kp} \quad \forall k, p \tag{9}$$

$$y_{ijpk}, x_k = 0, 1 \tag{10}$$

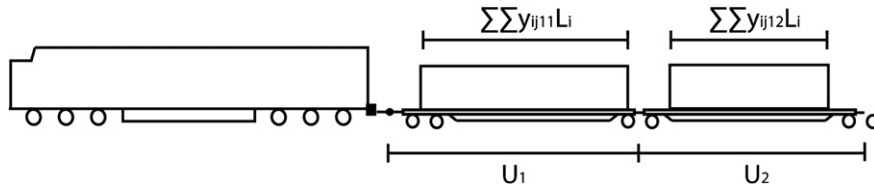


Fig. 4. Locomotive and first two intermodal units in a train.

The objective function (TAGL) represents the total adjusted gap length, which is comprised of two parts. The first part involves the gap length between the locomotive and the first load (Fig. 4), which is given by the difference between the length of the first unit (U_1) and the length of the load in position 1 of the 1st unit ($\sum \sum y_{ij1} L_i$) divided by 2. Multiplying the gap length by the adjusted factor A_1 results in the first adjusted gap length. The second part of the objective function computes the sum of the subsequent adjusted gap lengths. Each of the subsequent gaps is half of the difference in length between the current unit and the load ($(U_k - \sum \sum y_{ij1} L_i)/2$) plus half of the length difference between the next unit and the load ($(U_{k+1} - \sum \sum y_{ij1} L_i)/2$) multiplied by the appropriate adjusted factor, A_k . Note that we only take into account the loads in position 1 of all units in the train. This is reasonable since these are the only loads on spine or flat cars; and for well cars, the upper level gaps have a more significant aerodynamic effect than the lower level gaps (Furlong, 1988; Storm, 2005; Airflow Science Corporation, 2006).

Minimizing total adjusted gap length creates the most efficient train configuration. However, not all loads can be assigned to all slots. Loading assignment must conform to the loading capability of each unit as well as length and weight constraints. These are expressed in the model constraints above. The first two constraints, (3) and (4), ensure that each load is assigned to no more than one slot following the loading rules. Constraints (5) and (6) together state that if the bottom slot (position 2) in a well-car unit ($\delta_k = 1$) is not filled with loads greater than 40 ft, in which case Eq. (5) requires that $x_k = 0$, then no load can be assigned to the top slot (position 1) for the same unit, i.e., $\sum \sum y_{ij1k} = 0$ and therefore $y_{ij1k} = 0$ for all i, j . Constraint (7) ensures that containers cannot stack on top of trailers in well-car units. Constraints (8) and (9) are for weight and length limits, respectively. The weight constraint (8) is imposed for each car unit in order to reflect the total carrying capacity of that unit. The length constraint (9) is imposed for each slot to guarantee that the total length of loads in a given slot (position) does not exceed the length of that slot.

3.3. Solution algorithm

When assigning loads to IM trains, there are three possible scenarios that terminal managers may encounter: (1) number of loads = number of slots, (2) number of loads > number of slots, and (3) number of loads < number of slots. Scenario 1 and 2 are more common than scenario 3, and they can be solved directly by using the IP model to select the best loads for the available slots. When there are fewer loads than available slots (scenario 3), in some cases ALAM assigned two loads to two spine-car units (one load per unit) instead of double stacking them in a frontal well-car unit although a well-car unit was available. This is because the model only accounts for loads in position 1 (top) of all units for determining the adjusted gap length calculation. Clearly, such a loading pattern is not the most efficient alternative. To solve with this we developed an algorithm that can deal with all possible loading scenarios and successfully implement ALAM to determine the most energy efficient loading pattern.

A stepwise indirect approach is presented below based on the idea that placing loads towards the front of the train and leaving the rear cars empty is generally more aerodynamic and thus preferable. This approach ensures that only the front part of the train is first made available for loading; thereby leaving the rear of the train empty. The algorithmic details are provided below:

Step 1: $k = 1$, total number of loads = N_L , total number of units in the train = N_U

Step 2: Count number of slots from 1st unit to kth unit \rightarrow if this value is less than N_L and $k < N_U$ then go to Step 3; otherwise, go to Step 4

Step 3: $k = k + 1$ and go to Step 2

Step 4: Solve by ALAM \rightarrow if $k = N_U$ or there is no unassigned loads then go to Step 5; otherwise, go to Step 3

Step 5: Stop and output the optimal loading pattern

The algorithm starts with identifying the “loaded section” by increasing the number of available slots (in “loaded section”) until it is equal to the number of total loads. With the “loaded section” determined, we can then use ALAM to solve the IP problem. After implementing ALAM, if there are no unassigned loads or if k is equal to the total number of units in the train, we can output the optimal loading pattern; otherwise, we have to increase the number of available slots to accommodate the unassigned loads. Although the number of slots is equal to the number of loads, some loads might still be unassigned because of possible weight or length restrictions. If there are any unassigned loads and k is not the last unit in the train, the $(k + 1)$ st unit is made available.

With this solution procedure, ALAM is able to deal with all kinds of scenarios of IM loading which means the loading assignment model is now complete. This stepwise indirect algorithm may require several iterations before reaching the optimum; however, our computational experience shows that the solution time of ALAM is efficient enough, and therefore is a practical decision tool for real-time terminal operations.

4. Empirical application

Most IM trains operating in North America can be categorized into four general types: (I) International Stack Trains; (II) Domestic Stack Trains; (III) Trailer-on-Flat-Car (TOFC)/Container-on-Flat-Car (COFC) Trains; and, (IV) Mixed IM Trains consisting of both TOFC/COFC and Double Stack equipment (Armstrong, 1998). International and Domestic Stack Trains (Type I & II) have only well cars, TOFC/COFC Trains (Type III) have spine and flat cars, and mixed IM Trains (Type IV) are comprised of all types of IM railcars (well, spine, and flat cars).

We have developed a database of approximately 250 IM trains and loads that operated on the BNSF Chicago – Los Angeles route (BNSF Railway Company, 2005). Based on our assessment of their typical makeup, four representative trains were selected for detailed comparison of current terminal operations and the optimal loading pattern, given the loads and cars available (Table 2). Train 1 represents a double stack train transporting mostly international containers. Train 2 is also a double stack train but it is used for domestic containers. Train 3 is a TOFC/COFC train with a variety of trailers, and train 4 was comprised of well, spine and flat cars, and it also has a variety of differently-sized containers and trailers (Table 3).

As a result of the different characteristics of equipment and loads of these types of trains, there are varying degrees of flexibility in loading options. For example, train 1 has primarily 20' and 40' containers with little flexibility in alternative loading assignments compared to train 2, which had a much greater diversity of load

Table 2
Four representative IM trains

Train	Type	Number of cars	Number of units	Total loads
1	International stack train	30	104	224
2	Domestic stack train	37	115	244
3	TOFC/COFC train	37	110	131
4	Mixed train	32	104	173

Table 3
Number of loads in the example trains

Train	C20	C40	C45	C48	C53	T20	T28	T40	T45	T48	T53	Total loads
1	32	184	8	0	0	0	0	0	0	0	0	224
2	28	88	9	17	102	0	0	0	0	0	0	244
3	0	0	0	0	0	11	31	0	30	35	24	131
4	32	22	0	6	59	0	2	2	2	15	33	173

types due to the variation in domestic container size (Table 3). Consequently, the potential improvement in the aerodynamics of train 2 is greater than for train 1.

In the following sections, we present a series of scenario analyses for the four types of IM trains described above. In Section 4.1, we consider the scenario in which the number of loads equals the number of available slots in the train, and in Sections 4.2 and 4.3, we consider scenarios in which the number of loads is more, or less, than the number of slots, respectively. In Section 4.4, we compute fuel consumption for each of the scenario analyses, and the results and policy suggestions based on our analyses are presented in 4.5.

4.1. Scenario 1: number of loads equals number of slots

We used ALAM to analyze the optimal loading pattern for each of the four representative types of IM train and loads shown in Tables 2 and 3. Certain restrictions were applied to the loading pattern when assigning loads to slots (Association of American Railroads (AAR), 2004; Armstrong, 1998; TTX Company, 1999). For example, a 48'-well car cannot handle containers or trailers greater than 48' in the bottom position (position 2), although it can accommodate containers up to 53' in the top position (position 1). 53'-slot spine cars have only one platform (position 1) per unit and can handle either containers or trailers up to 53' long. Besides length and weight constraints, some of the IM units can accommodate only containers, some can accommodate only trailers, and the others can handle both. To ensure that load assignment follows the loading rules, possible loading combinations are specified for each IM unit. In order to clearly illustrate the aerodynamic effects, we assumed that none of the units were constrained by weight limits and the optimization process was based solely on minimization of the total adjusted gap length (TAGL).

For any given train and pool of loads, there is at least one loading pattern in which TAGL is minimized (the most aerodynamic pattern), and conversely, another loading pattern in which it is maximized (worst case). At present, terminal managers' goal in load assignment is to maximize slot utilization; therefore, they are largely indifferent to alternative loading patterns as long as they achieve 100% slot utilization (and comply with applicable loading rules). Consequently, with 100% slot utilization, average terminal operating practice will be some intermediate value between these two extremes. Therefore, we assumed that the mean of the TAGL for the minimum and maximum cases represents average terminal performance in the scenarios with 100% slot utilization (scenarios 1 and 2).

CPLEX 9.0 incorporated with GAMS (Brooke et al., 1998) was used to solve the model. Our computational experience showed that the optimal solution for each case could be obtained in less than five seconds. Thus, computational complexity is not a problem for real-time terminal operations. We then compared the results of ALAM with our estimate of average terminal loading (Fig. 5). In this scenario, the optimal results were generally close to the mean due to the inflexibility in possible loading combinations. This was especially evident for train 1 and train 3. Since train 1 has mostly international IM loads (i.e. 20' and 40' containers), the difference between the most and the least aerodynamic loading patterns is small. For train 3, the small range in TAGL is due to the characteristics of trailers and equipment in TOFC/COFC (Type III) trains. Train 4 is

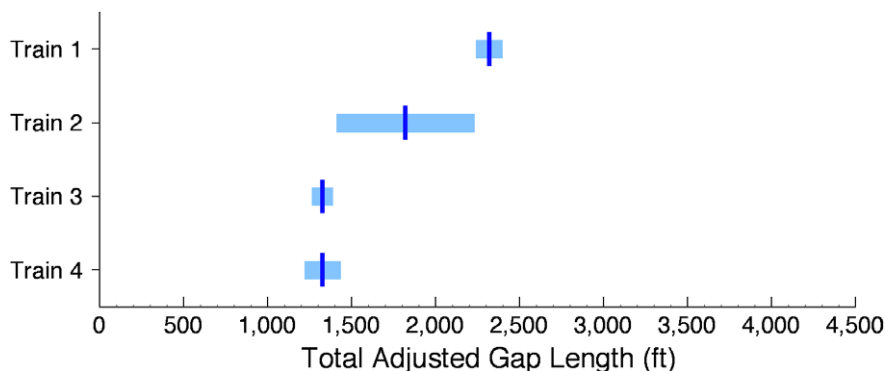


Fig. 5. Minimum, maximum and estimated average TAGL for the four example trains when loads = slots (Scenario 1).

Table 4
Number of available loads for example trains in scenario 2

Train	C20	C40	C45	C48	C53	T20	T28	T40	T45	T48	T53	Total loads
1	48	276	12	0	0	0	0	0	0	0	0	336
2	42	132	14	26	153	0	0	0	0	0	0	367
3	0	0	0	0	0	17	47	0	45	53	36	198
4	48	33	0	9	89	0	3	3	3	23	50	261

really a combination of trains with characteristics like trains 1 and 3 so it is not surprising that it also demonstrates a relatively small possible range of TAGL values for this scenario. By contrast, train 2 has a considerably greater TAGL range because of the wider variation in loads that is typical of this type of train.

4.2. Scenario 2: number of loads exceeds number of slots

At IM terminals, managers often have more loads than available slots in the next outgoing train (scenario 2) and this offers greater flexibility in load assignment. Scenario 2 is more common at IM terminals than scenarios 1 and 3 so these results are particularly important in assessing the potential benefits of using ALAM. The task in scenario 2 is to select the best set of loads from the load-pool to match the current outgoing train. Unfortunately, specific data on the composition of the pool of loads are not available because IM load assignments are only recorded when an outgoing train is loaded and ready to depart. However, it is reasonable to assume that the load pool is proportional to the distribution actually loaded onto a train. For the purpose of comparison among the four train types in this scenario analysis, we assumed that the number of each type of load was increased by approximately 50% (Table 4). Train consists and configurations were assumed to be unchanged.

Due to the greater number of potential loads to choose from compared to Scenario 1, there is some reduction in TAGL for all four train types; however, the magnitude of potential improvement varies widely (Fig. 6). Only train 1 (international double-stack container trains) experiences little potential improvement, again due to the inflexibility in types of loads. By contrast, all three of the other train types show substantial potential to reduce TAGL, with train 2 showing the greatest range. The wide range in possible TAGL means that terminal managers have a higher chance of loading aerodynamically inefficient trains and that use of ALAM has more potential to improve the energy efficiency of these types of trains.

4.3. Scenario 3: number of loads less than number of slots

In scenario 3, we analyzed the same four trains but the number of loads available was reduced by approximately 50% (Table 5). Since slot utilization is not 100% in this scenario, a larger adjusted gap length does not necessarily represent the poorest aerodynamics. According to Engdahl et al. (1987b), the worst aerodynamic

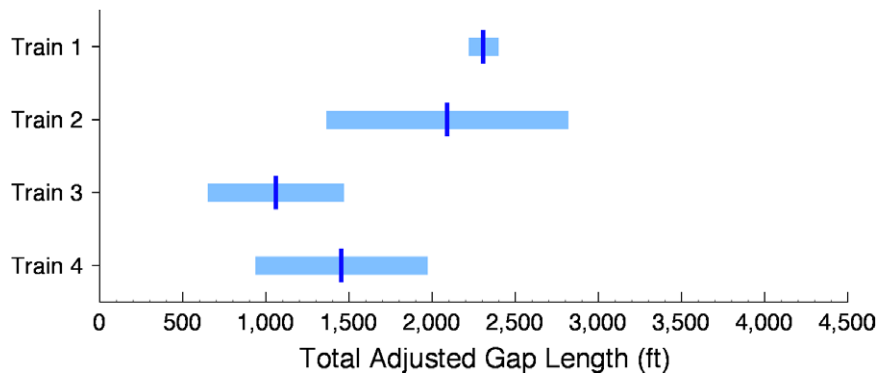


Fig. 6. Minimum, maximum and estimated average TAGL for the four example trains when loads > slots (Scenario 2).

Table 5
Number of available loads for example trains in scenario 3

Train	C20	C40	C45	C48	C53	T20	T28	T40	T45	T48	T53	Total loads
1	16	92	4	0	0	0	0	0	0	0	0	112
2	14	44	5	9	51	0	0	0	0	0	0	123
3	0	0	0	0	0	6	16	0	15	18	12	67
4	16	11	0	3	30	0	1	1	1	8	17	88

pattern is to have empty units uniformly distributed throughout the train. If terminal managers are loading without regard to aerodynamics, we expect loading patterns to be somewhere between the best and the worst. Therefore, we estimated their performance by assuming they placed half of the empty units at the end of the train, which would not affect aerodynamics, and the other half uniformly distributed in the “loaded section” (from the 1st unit to the last loaded unit).

The number of empty units for each of the four trains can be determined based on the optimal patterns obtained from ALAM (50, 57, 52, 38 units for trains 1, 2, 3, 4, respectively). For example, the best aerodynamic loading pattern for train 1 is to assign all available loads to the first 54 units (loaded section), and leave the last 50 units empty. The worst case for train 1 is to distribute the 50 empty units uniformly throughout the entire 104-unit train which results in 4446-ft adjusted gap length. Therefore, the average terminal performance was estimated by setting the last 25 units empty and distributing the other 25 empty units in the available loaded section. The differences in TAGL among optimal loading, worst loading and estimated average terminal performance are substantially greater than those in scenarios 1 and 2 because the calculation of adjusted gap lengths ignores units not in the loaded section (Fig. 7).

4.4. Fuel consumption computation

To quantify the potential fuel savings resulting from the optimal loadings obtained from ALAM, we computed the aerodynamic coefficients and the corresponding fuel consumption using the Aerodynamic Subroutine and the AAR Train Energy Model (TEM) (Furlong, 1988; Drish, 1992). The BNSF Transcon is a high speed freight rail route primarily with gentle grades, curves and rolling topography, so we selected a typical 100-mile segment to estimate fuel consumption and extrapolated this to develop estimates for the entire route.

Table 6 summarizes the computed fuel consumption values and associated savings if trains are loaded optimally using ALAM compared to estimated-average-loaded trains. The aerodynamic benefits of trains in scenario 2 are generally higher than scenario 1 due to the increased flexibility in loading patterns. However, there is almost no added benefit from optimizing loading train 1 compared to normal terminal practices because of the inflexibility in loading patterns for this type of train. Even though the flexibility of train 1 in scenario 2 is relatively higher than scenario 1, it still makes almost no difference in fuel consumption. We will further discuss the implication of this result in Section 4.5.2 below.

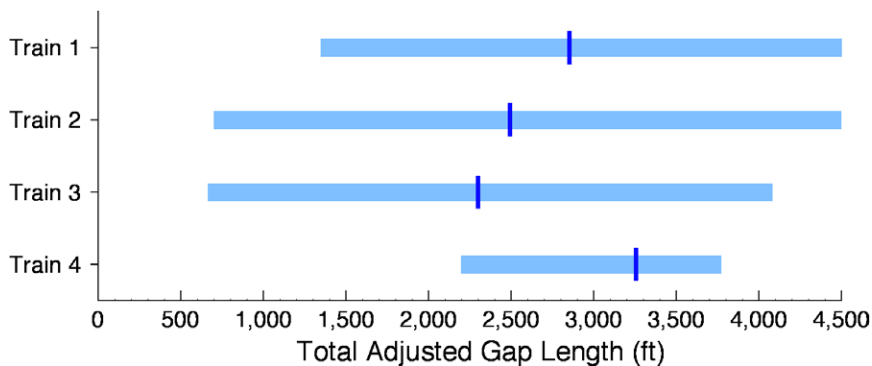


Fig. 7. Minimum, maximum and estimated average TAGL for the four example trains when loads < slots (Scenario 3).

Table 6

Fuel consumption and estimated savings for optimally loaded trains compared to average loaded trains (estimated terminal performance) in (a) scenario 1 (b) scenario 2 (c) scenario 3

Train	Fuel consumption (gallons)		Fuel savings	
	Optimal	Terminal	(gallons/mile)	%
(a)				
1	961	961	0.00	0.0
2	964	973	0.09	0.9
3	937	996	0.59	5.9
4	944	986	0.42	4.3
(b)				
1	961	961	0.00	0.0
2	955	988	0.33	3.3
3	937	1033	0.96	9.3
4	945	967	0.22	2.3
(c)				
1	827	992	1.65	16.6
2	825	1036	2.11	20.4
3	823	926	1.03	11.1
4	838	928	0.90	9.7

The aerodynamic benefit for train 3 (TOFC/COFC train) is generally higher than the others even though the range in TAGL in scenario 1 is not as large as the others (Fig. 5). This is because loads in spine or flat cars can be placed relatively closer to each other compared to the other types of trains comprised of well cars (Lai and Barkan, 2005), resulting in a greater difference in the aerodynamic efficiency of loading patterns.

Scenario 3 is a relatively unique case (slot utilization \ll 100%) compared to scenarios 1 and 2 (slot utilization = 100%). The differences between optimal patterns and estimated average terminal practices in scenario 3 are substantial because there are quite a few long gaps in the terminal case caused by empty slots in the train (Table 6c). The fuel savings for trains 1 and 2 are higher here since the aerodynamic drag caused by empty slots between two double-stack units is greater than empty slots between single-level units in trains 3 and 4.

4.5. Policy recommendations for railway intermodal operations

The scenario analyses presented above demonstrate the potential benefit of implementing ALAM at IM terminals. In this section, we discuss the implications of these results and some options to improve the energy efficiency of IM train operations.

4.5.1. Use of ALAM for intermodal train loading

The most obvious recommendation stemming from the work described here is that ALAM be integrated into railroads' IM terminal loading operations. Use of ALAM to match loads to slots would have reduced fuel consumption by as much as 20%. The exact amount will vary for individual train consists and loads available, but these figures provide insight into the magnitude of potential savings available. The percentage of each of the four types of trains on the BNSF Transcon are approximately equal with a total of about 50 per day. Accounting for the different potential savings for these train types if they were loaded according to ALAM guidelines, compared to the average loading configuration, translates to a total annual savings of about 28 million dollars based on the most common scenario (scenario 2).

Major railroad IM terminals already use software to assist and expedite the loading process, so integration of ALAM objectives and methodology into this software would not require substantial institutional or process change. It also should have little if any impact on operating cost because it will not generally require more work, but rather, inform the loading process with a new quantitative parameter, minimization of TAGL. Implementation of ALAM would automate terminal managers' consideration of the large variety of loads and railcar types available thereby enabling them to load trains in a more aerodynamically efficient configuration.

These benefits can be further enhanced by several additional factors as described in the following three subsections.

4.5.2. *Better matching of loads with railcars*

The aerodynamic benefit of optimizing the loading pattern of train 1 in scenarios 1 and 2 is small. This is due to the inflexibility in loading patterns of train 1, and also the characteristics of the “long” well cars. Train 1 is primarily transporting 20’ and 40’ international containers with just a few 45’ containers. Nevertheless, approximately 70% of its railcars were designed for 53’ containers. Placing 45’ containers in the frontal positions in the train would generally be better than 40’ containers; however, for this kind of long well-car unit, it makes little aerodynamic difference. This is because using 53’ cars to transport either 40’ or 45’ loads always results in gaps greater than the critical gap length (12 ft), necessary to gain aerodynamic benefit (Engdahl et al., 1987a) so reconfiguring the shorter loads on the longer cars has little effect (Lai and Barkan, 2005).

Using 53’-well cars is convenient for managers because of these cars’ flexibility; however, placing shorter international containers in 53’-well cars causes greater aerodynamic resistance than if 40’-unit well cars are used. For example, without changing the placement of loads, if train 1 used well cars designed for 40’ containers (45’ containers can still be placed in at the top), the aerodynamic coefficient reduces from 7.04 to 5.50 lbs/mph², and the weight of the train would be reduced by 18% as well. The corresponding fuel savings would be 0.88 gallons per mile, an approximately 9% reduction. Consequently, we suggest better matching IM loads with the railcars used to transport them, specifically acquisition and use of well cars with 40’ slots that are designed for international loads. In recognition of the growth in international container traffic and consistent with this recommendation, one railroad has a car modification program converting some existing 48’-well cars into 40’-well cars (Stehly, 2007).

4.5.3. *Optimize loading for more than a single train simultaneously*

The analysis thus far has focused on optimizing the aerodynamic efficiency of a single outgoing train for a given set of loads. We formulated the problem this way because most loads arrive shortly before loading begins. However, if advance knowledge (either empirical or probabilistic) on the composition of outgoing trains and the load pool is available, this information can be used to optimize the loading of multiple trains simultaneously and increase the benefit of applying the model developed in this paper.

ALAM can be extended to address the multiple train loading problem by modifying it to consider units of all the loads and units in each of the trains available. This is accomplished by introducing a new index, *t*, that refers to the set of outgoing trains. The objective function of the extended model is minimization of the total adjusted gap length of all the outgoing trains in a given time horizon:

$$\text{Min } \sum_{t=1}^T z_t \tag{11}$$

Where:

$$z_t = 0.5 \times \left\{ A_1 \left(U_{t1} - \sum_i \sum_j y_{ij1} L_i \right) + \sum_{k=1}^N A_{k+1} \left[\left(U_{tk} - \sum_i \sum_j y_{ij1k} L_i \right) + \left(U_{tk+1} - \sum_i \sum_j y_{ij1k+1} L_i \right) \right] \right\}.$$

The constraints are the same as ALAM except the decision variables y_{ijpk} and x_k are replaced by y_{ijtpk} and x_{tk} . Compared to the basic ALAM described earlier, this modified version provides a global optimum solution as opposed to the local optimum based on analysis of a single train.

4.5.4. *Uncouple empty railcars at the end of IM trains*

The optimal loading pattern aims to achieve the lowest aerodynamic resistance and thereby maximize fuel economy; therefore, when the number of loads is less than the number of slots (scenario 3), the loads are placed in cars at the front of the train, and empty slots in the rear. According to IM loading rules, the speed of a train with empty cars is restricted to 55 mph due to concerns about dynamic instability of these cars at higher speeds (BNSF Railway Company, 2004b). Hence, there is a tradeoff between aerodynamic loading

Table 7

Fuel consumption and estimated savings for the worst case, average loaded trains (estimated terminal performance), and optimal results without or with uncoupling policies in scenario 3

Train	Fuel consumption (gallons)				Fuel savings	
	Worst case	Terminal	Optimal without uncoupling	Optimal with uncoupling	(gallons/mile)	%
1	1090	992	827	598	3.94	39.7
2	1135	1036	825	562	4.74	45.8
3	976	926	823	578	3.48	37.6
4	976	928	838	706	2.22	23.9

pattern and train operating speed. Although ALAM is intended to maximize aerodynamic efficiency, it can be modified to suit terminal operators' or dispatchers' preferences (higher train speed vs. better fuel efficiency) by adding additional constraints or pre-processes (such as each car must have at least one load so that the speed restrictions can be avoided).

We can also approach this problem by removing the empty cars from the end of the train. Although this alternative might incur some additional operating costs, it should be compared to the energy efficiency benefits. Uncoupling empty cars would reduce the weight of the train and also eliminate the speed restriction. Table 7 shows the computed fuel consumption and respective savings for the worst case, average terminal practices, and optimal loading pattern with or without an uncoupling policy. Fuel savings are computed by comparing terminal practices to the optimal results if the uncoupling strategy is used. The savings with implementation of the practice of uncoupling are greater than without uncoupling. Comparison of Table 6c and Table 7 shows that the uncoupling strategy can increase fuel savings by 14–25%. Therefore, if repositioning empty IM equipment can be efficiently accomplished without transporting it in IM trains with loads, uncoupling railcars from the end of trains should be considered.

5. Discussion and future research

ALAM can be extended to optimize the aerodynamic efficiency at the system level instead of optimizing the loading of a single train. The time horizon can be extended to several days (or more if desired) depending on the availability of advance train-and-load information and how difficult it is to solve the problem using IP. An interesting question we intend to address in the future is to consider the time horizon required to approach the global optimum, and how much improvement is possible as that horizon is lengthened.

This study focused on IM services between Chicago and LA in which no intermediate train reconfiguration occurs en route. This is the case for approximately 80% of the trains operating on this route. In the majority of cases, the initial loading pattern could be determined to optimize the train's aerodynamic efficiency without considering unloading sequences. Of the remaining 20%, half of them have no more than two additional terminals and these are generally near the final destination so the same aerodynamic benefits will apply over most of the route. If intermediate operations do occur, then the aerodynamic benefits will apply for the portion of the route prior to the terminal where the train is reconfigured. Beyond that trains may operate at lower efficiency. It would also be possible to modify the model to optimize aerodynamic efficiency for the entire route including intermediate destinations (where the train is reconfigured) in the analysis. However, there is a trade off between the costs and benefits of doing this. The increased modeling complexity would demand more computation time and the resulting problem might be too complex to solve using IP. Further research should be conducted to develop specially designed solution algorithms that can generate very good (if not optimal) loading patterns for trains whose composition changes en route.

6. Conclusions

We develop a mathematical programming model (ALAM) by incorporating the aerodynamic characteristics of intermodal trains to optimize load-to-unit assignments. The model can be integrated into current terminal operation software as an additional tool to help terminal managers make better loading decisions. The

contributions of this paper to the literature are: (1) It is the first use of optimization modeling with the objective of improving the aerodynamics and consequent energy efficiency of intermodal trains and reveals significant possible savings are possible. (2) The model developed in this paper can be adapted to a variety of other intermodal train loading assignment problems through modification of the objective function. This is a novel contribution to the literature and enhances its generality because the formulation can be solved efficiently and thus serve as a basis for other intermodal load assignment problems. (3) Several policy recommendations regarding railway intermodal operations are developed based on a series of scenario analyses.

There are substantial potential fuel and cost savings benefits that railroads can achieve through implementation of ALAM at intermodal terminals. These benefits can be further enhanced through several additional steps including: (a) better matching of railcars and loads for international intermodal trains (b) simultaneous optimization of multiple trains to take greater advantage of the potential to improve energy efficiency of intermodal trains through use of more aerodynamic loading patterns, and (c) uncoupling empty railcars from the end of loaded intermodal trains when practical. The potential annual savings in fuel consumption through use of ALAM by one large railroad on one of its major intermodal routes is estimated to be approximately 15 million gallons with a corresponding value of 28 million dollars. Correspondingly larger savings in fuel, emissions and expense are possible if the methodology described in this paper were applied to all North American intermodal trains.

Acknowledgements

The authors are grateful to Mark Stehly, Larry Milhon, Paul Gabler, and Mark Dinger for their assistance on this project. This research was supported by a grant from the BNSF Railway Technical Research and Development Program. A portion of the first author's graduate study support has been from a CN Railroad Engineering Research Fellowship at the University of Illinois at Urbana-Champaign.

References

- Airflow Science Corporation, 2006. Aerodynamic Subroutine Version 4.0. Association of American Railroads, Washington, DC.
- American Railway Engineering and Maintenance-of-Way Association (AREMA), 2001. Manual for Railway Engineering, Chapter 16, Part 2, Train performance. American Railway Engineering and Maintenance-of-Way Association, Landover, MD.
- Armstrong, J., 1998. *The Railroad: What It Is, What It Does*, fourth ed. Simmons-Boardman Books, Inc., Omaha, NE.
- Association of American Railroads (AAR), 1981. Proceedings Railroad Energy Technology: The Alternatives. Railway Systems & Management Association, North Field, NJ.
- Association of American Railroads (AAR), 1987. Proceedings Railroad Energy Technology Conference II. Association of American Railroads, Chicago, IL.
- Association of American Railroads (AAR), 2004. Loading Capabilities Guide. Association of American Railroads, Washington, DC.
- Association of American Railroads (AAR), 2004–2006 editions. Railroad Facts. Association of American Railroads, Washington, DC.
- Bertin, J.J., 2001. *Aerodynamics for Engineers*. Prentice Hall, NJ, pp. 85–88.
- BNSF Railway Company, 2004a. Every drop counts in goal to improve fuel efficiency. BNSF News, BNSF Railway Company, Fort Worth, TX.
- BNSF Railway Company, 2004b. Intermodal Loading Guide. BNSF Railway Company, Fort Worth, TX.
- BNSF Railway Company, 2005. Train Documents. BNSF Railway Company, Fort Worth, TX.
- Brooke, A., Kendrick, D., Meeraus, A., Raman, R., 1998. GAMS – A User's Guide. GAMS Development Corporation, Washington, DC.
- Burriss, C., 2003. Eastbound Statistical Analysis: Slot Utilization & Mixed Cars. OASIS Users' Conference, Austin, TX.
- Corry, P.G., Kozan, E., 2006. An assignment model for dynamic load planning of intermodal trains. *Computers and Operations Research* 33, 1–17.
- Drish, W.F., 1992. Train Energy Model Version 2.0 Technical Manual. SD-485. Association of American Railroads, Chicago, IL.
- Engdahl, R., Gielow, R.L., Paul, J.C., 1987a. Train resistance – aerodynamics volume I of II intermodal car application. In: Proceedings of Railroad Energy Technology Conference II. Association of American Railroads, Chicago, IL.
- Engdahl, R., Gielow, R.L., Paul, J.C., 1987b. Train resistance – aerodynamics volume II of II open top car application. In: Proceedings of Railroad Energy Technology Conference II. Association of American Railroads, Chicago, IL.
- Engdahl, R., 1987. Full-Scale Rail Car Testing to Determine the Effect of Position-in-Train on Aerodynamic Resistance. SD-705. Association of American Railroads, Washington, DC.
- Feo, T.A., Gonzalez-Velrade, J.L., 1995. The intermodal trailer assignment problem. *Transportation Science* 29, 330–341.
- Furlong, C.F., 1988. Aerodynamic Subroutine Users Guide. SD-683. Association of American Railroads, Washington, DC.
- Gallamore, R.E., 1998. State of the art of intermodal freight transport in the United States. In: *Intermodal Freight Transport in Europe and the United States*. Eno Transportation Foundation, Inc., Lansdowne, VA.

- Gielow, M.A., Furlong, C.F., 1988. Results of Wind Tunnel and Full-Scale Tests Conducted from 1983 to 1987 in Support of the Association of American Railroads' Train Energy Program. SD-685. Association of American Railroads, Washington, DC.
- Hay, W.W., 1982. *Railroad Engineering*, second ed. John Wiley & Sons, Inc., New York, NY.
- International Union of Railways (UIC), 2004. Proceedings 2nd UIC Railway Energy Efficiency Conference. International Union of Railways, Paris.
- Lai, Y.C., Ahuja, N., Barkan, C.P.L., Drapa, J., Hart, J.M., Jawahar, C.V., Kumar, A., Milhon, L., Narayanan, P.J., Stehly, M., 2007. Machine vision analysis of the energy efficiency of intermodal freight trains. *Journal of Rail and Rapid Transit* 221, 353–364.
- Lai, Y.C., Barkan, C.P.L., 2005. Options for improving the energy efficiency of intermodal freight trains. *Transportation Research Record* 1916, 47–55.
- Muller, G., 1999. *Intermodal Freight Transportation*, fourth ed. Eno Transportation Foundation and Intermodal Association of North America, Washington, DC.
- Optimization Alternatives Ltd. Inc. Optimization Alternatives' Strategic Intermodal Scheduler (OASIS). <<http://www.oax.com>> Accessed 06.01.07.
- Powell, W.B., Carvalho, T.A., 1998. Real-time optimization of containers and flatcars for intermodal operations. *Transportation Science* 32, 110–126.
- Smith, M.E., 1987. Economics of reducing train resistance. In: Proceedings of Railroad Energy Technology Conference II. Association of American Railroads, Chicago, IL.
- Stehly, M., 2007. Personal Communication. BNSF Railway Company, Fort Worth, TX.
- Stodolski, F., 2002. Railroad and Locomotive Technology Roadmap. ANL/ESD/02-6. Argonne National Laboratory, Argonne, IL. <<http://www.climatevision.gov/sectors/railroads/pdfs/roadmap.pdf>> Accessed 08.26.07.
- Storm, B., 2005. Personal Communication. NASA Ames Research Center, Moffett Field, CA.
- TTX Company, 1999. Equipment Guide. TTX Company, Chicago, IL.
- Union Pacific Railroad Corporation, 2004. Intermodal Loading Guide. <<http://www.uprr.com/customers/dam-prev/attachments/intgenloadguide.pdf>> Accessed 06.01.07.
- Utterback, M., 2006. Personal Communication. BNSF Railway Company, Chicago, IL.