F. Piu, V. Prem Kumar, M. Bierlaire, M. G. Speranza

Exploiting Homogeneity Aspects for Locomotive Scheduling Problems

WPDEM 2013/13
Exploiting Homogeneity Aspects for Locomotive Scheduling Problems

F. Piu \(^{(1)}\) \hspace{1em} V. Prem Kumar \(^{(2)}\) \hspace{1em} M. Bierlaire \(^{(3)}\) \hspace{1em} M. G. Speranza \(^{(4)}\)

\(^{(1)}\) Department of Economics and Management, University of Brescia, C. da S. Chiara 50, Brescia, Italy, francesco.piu@eco.unibs.it

\(^{(2)}\) TRANSP-OR laboratory, École Polytechnique Fédérale de Lausanne, Station 18, Lausanne, Switzerland, prem.viswanathan@epfl.ch

\(^{(3)}\) TRANSP-OR laboratory, École Polytechnique Fédérale de Lausanne, Station 18, Lausanne, Switzerland, michel.bierlaire@epfl.ch

\(^{(4)}\) Department of Economics and Management, University of Brescia, C. da S. Chiara 50, Brescia, Italy, speranza@eco.unibs.it

Abstract

The \textit{Locomotive Assignment Problem} (LAP) is a class of planning and scheduling problems solved by assigning a fleet of locomotives to a network of trains, optimizing one or more objectives and satisfying a rich set of constraints. In the planning versions of the LAP, we determine the type of \textit{consist} (a group of linked locomotives) assigned to each train in a given schedule. The set of the consist types that are initially available is assumed to be given in the prior works, while the planning versions of the LAP determine the number of consist units used for each consist type. We introduce an optimization model (called \textit{consists selection}) that precedes the planning LAP solution and determines the set of consist types. The consist types are selected not only by minimizing the active and ownership costs, but also by the number of fueling events and the inefficiencies in the fuel capacity exploitation. This selection leads to solutions that produce savings in terms of overall fueling cost and are easier to handle in the routing phase. We demonstrate yearly savings up to 201 fueling events (985 servicing hours, US$ 117000) for a set of 229 trains by solving a series of realistic problem instances.

\textbf{Keywords:} locomotive assignment, multicommodity flow, planning, scheduling
1 Introduction

The Locomotive Assignment Problem (LAP) is a class of planning and scheduling problems among the most complex and important in the railway sector because it involves a large number of expensive assets. The LAP is solved assigning a fleet of locomotives to a network of trains optimizing one or more crucial objectives (costs, profit, fleet size, level of service) and satisfying a rich set of technical and budget constraints. Large-scale very complex freight rail activities impose to separate the LAP in three distinct problems:

1. locomotive planning problem (LPP);
2. locomotive scheduling problem (LSP);
3. locomotive routing problem (LRP).

The LPP, the LSP and the LRP are solved sequentially. At the beginning of the solution sequence (in the LPP) we consider the available locomotive types while at the end (in the LRP) we route specific locomotive units to fueling and maintenance stations imposing the required fueling and maintenance stops (i.e. honoring the fueling and maintenance constraints). The separation of the LAP leads to definitely suboptimal solutions. To achieve optimality a model should encompass LPP, LSP and LRP. Such a structural integration is not practicable due to the size and complexity of real problems. This study focuses on the planning versions of the LAP (i.e. the LPP) and on its integration with the LRP. A consist is a combination of locomotives and the types of consist are represented by the combinations of locomotive types. In the LPP for each train we determine the type of consist assigned to that train. The set $C$ of the consist types that are initially available to solve the LPP is generally taken as given in terms of consist types (railroad companies provide $C$ to researchers). Given $C$, the LPP solution determines the number of consist units used for each consist type. This research does not take $C$ as given. Instead, we introduce an integer linear programming model (called consist types selection or shortly consists selection) to identify the consist types included in $C$. The consists selection precedes the LPP solution and determines the consist types that form $C$ and are available to solve the LPP. We select the consist types by minimizing the active and ownership costs but also the number of fueling events and the inefficiencies in the consist fuel capacity exploitation. This selection leads to LPP solutions that produce savings in terms of overall fueling cost and are easier to handle in the routing phase. This methodological innovation partially integrates planning and routing phases accounting (indirectly) for the fueling constraints in the LPP. Solving several realistic instances we obtain yearly savings up to 201 fueling events (985 servicing hours, US$ 117000 current value) for a set of 229 trains.
2 The Locomotive Planning Problem

North American freight trains are generally very heavy and a single locomotive is often not sufficient to satisfy the required motive power performance. Therefore a suitable consist has to be chosen from a set of locomotives with different characteristics. The solution of the LPP determines, for each train, the type and the number of locomotives assigned to that train. Usually, in the LPP the train schedule is given and cannot change (delays and disruptions are excluded). In the LPP only the number of locomotives and their type matter, the specific identification number of each locomotive is not considered and locomotives of the same type are completely equivalent (while in the LSP and LRP we work with specific locomotive units considering their unique locomotive identification numbers).

2.1 Definitions

Before continuing, we provide a more precise characterization of locomotives and trains and we introduce some definitions.

A locomotive is characterized by its:

1. maximum Horse Power (HP);
2. maximum pulling force or Tractive Effort (TE);
3. range (i.e. fuel tank capacity and fuel consumption rate).

The term train indicates a train service characterized by:

1. \(\langle \text{departure time, departure station} \rangle\) and \(\langle \text{arrival time, arrival station} \rangle\).
2. TE requirement (depends on train weight and track geometry).
3. HP requirement (imposed by train speed).

Active locomotives pull trains but locomotives may also move in a passive way. Deadheading locomotives may be attached to trains as passive rolling stock elements and are moved like wagons in order to be repositioned. Light-travelling locomotives may form a group where only the leading locomotive is active and pulls the remaining locomotives attached as passive rolling stock elements. Freight trains may be classified in fast and slow trains. According to AREMA [2003], HP and TE are related via the maximum speed achievable.
by a train, namely speed is proportional to $\frac{HP}{TE}$. Consequently, high HP consists are suitable for fast freight trains (intermodal and auto trains) while low HP consists are suitable for slow freight trains (merchandise and bulk trains). Since the reduction in operations costs is primarily pursued by minimizing the number of used locomotives, it is important to promote train to train connections: when a train service ends at its arrival station (say station $\mathcal{S}$) its consist is assigned to a compatible outbound train (whose departure station is $\mathcal{S}$) in its entirety. The reassignment of a consist in its entirety would avoid consist busting operations, where the consist busting is the operation of merging locomotives from inbound trains and regrouping them to make new consists to be assigned to outbound trains. The consist busting typically entails the breaking up of an incoming consist at a station and the assignment of the locomotives in it to more than one outgoing train. According to Vaidyanathan et al. [2008a], consist bustings are characterized by labor, cost and time intensive activities (each consist busting requires between two to six additional hours per locomotive within the station). The ownership and the active utilization of a consist have costs that are expressed in US dollars and are specific for each consist type. A primary part of the information sources exploited in this research reports several locomotive and train costs in terms of US$ in 2008 ($2008US\$$). For this reason we have expressed all the monetary values in terms of 2008US$ throughout this paper.

3 Literature review

The LAP may be classified into several categories depending on the adopted classification scheme. A possibility is to classify the LAP looking at the application field:

1. freight trains;
2. passenger trains;
3. industrial in-plants railroads;
4. switch locomotives (yard switching).

Few studies focus on the last two version of the LAP, see for example Lübbecke and Zimmermann [2003] for the in-plant railroad LAP and Sabino et al. [2010] for the switch LAP. We found few works about the management of passenger trains locomotives reported in the operations research literature (Cordeau et al. [1998] present a survey on this subject), while more papers focus on the freight railway locomotives assignment (see Piu and Speranza [2012] for a recent survey on the LAP). There are significant differences in
modeling and complexity of the planning LAP in the passenger and freight framework. Quite often a single locomotive is sufficient to pull a passenger train and when more than one locomotive is needed, the consist is usually constituted by no more than two locomotives of the same type. The reduced size of passenger trains and consists make the problem more tractable with respect to the freight version. Thereby, it is possible to assign simultaneously both locomotives and cars to passenger trains (Cordeau et al. [2000], Cordeau et al. [2001], Lingaya et al. [2002]), while for freight trains these two assignments are managed separately. Tables I and II summarize our findings and report a list of researches inspired by real problems.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Railway company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florian et al. [1976]</td>
<td>Canadian National</td>
</tr>
<tr>
<td>Ziarati et al. [1997]</td>
<td>Canadian National</td>
</tr>
<tr>
<td>Ziarati et al. [1999]</td>
<td>Canadian National</td>
</tr>
<tr>
<td>Ziarati et al. [2005]</td>
<td>Canadian National</td>
</tr>
<tr>
<td>Ireland et al. [2004]</td>
<td>Canadian Pacific Railway</td>
</tr>
<tr>
<td>Ahuja et al. [2005a]</td>
<td>CSX Transportation</td>
</tr>
<tr>
<td>Ahuja et al. [2005b]</td>
<td>CSX Transportation</td>
</tr>
<tr>
<td>Ahuja et al. [2006]</td>
<td>CSX Transportation</td>
</tr>
<tr>
<td>Vaidyanathan et al. [2008a]</td>
<td>CSX Transportation</td>
</tr>
<tr>
<td>Vaidyanathan et al. [2008b]</td>
<td>CSX Transportation</td>
</tr>
<tr>
<td>Powell and Bouzaïene-Ayari [2007]</td>
<td>Norfolk Southern</td>
</tr>
<tr>
<td>Brannlund et al. [1998]</td>
<td>Banverket Swedish National Railway</td>
</tr>
<tr>
<td>Scholz [2000]</td>
<td>Stateiis Järivägar Swedish State Railways</td>
</tr>
<tr>
<td>Noble et al. [2001]</td>
<td>Public Transport Corporation</td>
</tr>
<tr>
<td>Baceler and Garcia [2006]</td>
<td>Companhia Vale do Rio Doce</td>
</tr>
<tr>
<td>Fügenschuh et al. [2006]</td>
<td>Deutsche Bahn AG</td>
</tr>
<tr>
<td>Fügenschuh et al. [2008]</td>
<td>Deutsche Bahn AG</td>
</tr>
</tbody>
</table>

Table I: Freight railway locomotives management.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Railway company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramani [1981]</td>
<td>Indian Railways</td>
</tr>
<tr>
<td>Cordeau et al. [2000]</td>
<td>VIA Rail Canada</td>
</tr>
<tr>
<td>Cordeau et al. [2001]</td>
<td>VIA Rail Canada</td>
</tr>
<tr>
<td>Lingaya et al. [2002]</td>
<td>VIA Rail Canada</td>
</tr>
<tr>
<td>Illés et al. [2005]</td>
<td>Magyar Államvasutak</td>
</tr>
<tr>
<td>Illés et al. [2006]</td>
<td>Magyar Államvasutak</td>
</tr>
<tr>
<td>Paoletti and Cappelletti [2007]</td>
<td>Trenitalia</td>
</tr>
</tbody>
</table>

Table II: Passenger railway locomotives management.
From a modeling perspective, the first important classification is obtained considering the maximum number of pulling locomotives a train may require. If each train needs a single pulling locomotive, then the problem is modeled by a *single locomotive model* (see for example Forbes et al. [1991] and Fügenschuh et al. [2006]). If some trains require more than one pulling locomotive then the problem is modeled by a *multiple locomotive model*.

### 3.1 Multiple locomotive models

The most complex version of the LAP is solved when trains are pulled by consists obtained joining several locomotives of different types.

Florian et al. [1976] analyzed a problem faced by the Canadian National Railways (CN) and were among the first to deal with this version of the LAP. They formulated the problem as an integer program based on a multi-commodity network flow formulation. The objective was to minimize the capital investment and the maintenance costs over a long planning horizon.

Ziarati et al. [1997] extended the formulation proposed in Florian et al. [1976] to include many of the operational constraints encountered at CN. Ziarati et al. [1997] proposed a space-time network approach for the operational version of the LAP and formulated the problem as a mixed integer linear programming (MILP) problem corresponding to a multi-commodity network flow problem with supplementary variables and constraints.

Ahuja et al. [2005b] significantly improved the realism of the LPP models studying a weekly locomotive scheduling problem faced by CSX Transportation. Ahuja et al. [2005b] proposed a mixed integer programming (MIP) formulation and modelled the problem as a multicommodity flow (each locomotive type corresponds to a different commodity) with side constraints (the number of locomotives of each type is limited) on a space-time network where arcs denote trains, and nodes denote events, i.e. arrivals and departures of trains and locomotives. Having defined the total cost as the sum of ownership, active, deadheading, light-traveling and consist busting costs plus the penalty for the use of single-locomotive consists, the objective is to minimize the overall costs identifying active, deadheading, light-traveling locomotives and train-to-train connections. A technical document (Ahuja et al. [2006]) indicates some possible extensions of the model like optimal routing of locomotive to fueling stations and shops to satisfy fueling and maintenance constraints.

Vaidyanathan et al. [2008a] implement part of these extensions and in particular propose the *Consist Flow Formulation* for the LPP.
The LPP may be solved starting from a set of locomotive types not already assembled in consists or may be solved starting from a set \( C \) of consist types fixed ex-ante. The *Locomotive Flow Formulation* (LFF) described in Ahuja et al. [2005b] defines each locomotive type as a commodity while the *Consist Flow Formulation* (CFF) replaces locomotive types with consist types and each consist type is defined to be a single commodity. In the LFF, single locomotives are assigned to trains, and the consists are the result of these assignments while in the CFF the solution is obtained starting from a set of already assembled consists. As expected, the solution quality critically depends on the number and types of available consists: a greater number of consist types leads to a higher solution quality. The use of assembled consists restricts the solution space leading to a loss in solution quality. Nevertheless, computational tests performed by Vaidyanathan et al. [2008a] show that the optimal objective function value in the CFF may be just 5% higher than the one obtained in the LFF. The correct identification of the set of assembled consists is crucial to reduce as much as possible the optimality gap. This (small) optimality gap is highly compensated by many benefits:

1. in several instances the LFF may not converge to a feasible solution in more than 10 hours, while the CFF optimally solve the same instances within a few minutes;

2. the CFF allows to implicitly handle many constraints that were explicitly used in the LFF, offering shorter computation and rapid convergence;

3. complex rules on the allowed consist classes (locomotive types combinations), impossible to impose in the LFF, are easy to enforce in the CFF;

4. consist busting (and its corresponding cost) is reduced to a large extent.

As some important real-life constraints cannot be inserted in the LPP, the models proposed in Ahuja et al. [2005b] and Vaidyanathan et al. [2008a] did not account for the fueling and maintenance constraints that are honored in the locomotive routing phase.

4 **The current study**

The literature survey shows the lack of studies focused on the integration of planning, scheduling and routing phases. The three phases are solved sequentially: the solution of the LPP is the starting point for the subsequent scheduling and routing problems. Important constraints (like fueling and maintenance constraints) are
relegated in the routing phase because they rely on locomotives identification numbers that are not considered in the LPP.

This study proposes a methodological innovation able to partially integrate planning and routing phases accounting (indirectly) for the fueling constraints in the consist flow formulation of the LPP (CFF of the LPP). In the previous studies, the set \( C \) of available consist types were assumed as given. This means that the consist types that are available for the LPP optimization are determined by the expertise of locomotive managers. Subsequently, the LPP solution determines the number of consist units needed for each consist type in \( C \).

Due to the high number of possible consist types, it is impracticable to adopt a set \( C \) obtained through an explicit enumeration of all the possible consist types.

Our objective is to define a preliminary optimization program (called *consist types selection* or shortly *consists selection*) that determines the set \( C \) identifying the consist types initially available for the solution of the LPP. The consists selection indirectly accounts for a reduction of the number of fueling events. This phase may identify consist types that are not captured by a purely cost-oriented consists selection but that may be useful in the routing phase, where a reduced number of fueling events may simplify fueling routing and produce savings that could not be achieved using (apparently) more economical consist types. Finally, we exploit in the consists selection the concept of consist fueling homogeneity that allows the identification of consist types that efficiently exploit their fuel capacity and reduce the fueling costs.

5 Consists selection: concepts and methodology

Inspired by the CFF of the LPP proposed in Vaidyanathan et al. [2008a], we introduce an optimization program that precedes the LPP optimization and selects the initially available consist types among all the possible consist types.

The aim of the consists selection is to identify the composition of the consist types set \( C \) i.e. to define the consist types initially available for the solution of the LPP. The actual number of units for each consist type will be precisely determined after solving the LPP. Our objective is to identify a set \( C \) that gives LPP solutions easier to handle when fueling constraints are satisfied. These solutions reduce the number of fueling stops in the routing phase and may turn out to be more economical when we consider the planning and the routing phases altogether.
5.1 The actual number of fueling stops

To reduce the total fueling cost we may diminish the fueling stops cost (FSC) by reducing the total number of fueling events. Given some operative conditions, each locomotive type is characterized by its fuel consumption rate. Given the consumption rate, the actual number of fueling stops is greater than the one calculated relying on the fuel nominal tank capacity (hereinafter the term tank capacity refers to the actually exploitable fuel tank capacity while the term nominal tank capacity refers to the nominal volume of the tank). According to Ahuja et al. [2006] railroads have been found to have several out of fuel events in a day. Out of fuel events have severe costs, US$ 8000 each in 2000 according to GE Harris Energy Systems [2000] (equivalent to 2008US$ 9995). These costs could be avoided by measuring electronically the fuel level. However, according to Lindsey [2007], in 2008 nearly 90 percent of the locomotives were still without electronic fuel measurement. Lindsey [2007] reports that railroads adopt very conservative practices to avoid out of fuel events, information confirmed by GE Harris Energy Systems [2000] (locomotives are refueled when their fuel tanks are still 60% full) and by Ahuja et al. [2006] (the average fuel dispensed per event is only one third to half the locomotive tank capacity). The range of each locomotive type depends on its consumption rate and on its fuel tank capacity. These very conservative policies reduce the actually exploited fuel tank capacity increasing the number of fueling events (and the overall fueling cost). According to GE Harris Energy Systems [2000], we assume for each locomotive type an actually exploited fuel tank capacity equal to 40% of the nominal tank capacity.

5.2 The fueling stop cost

According to Unkle and Roddy [2004], a locomotive may be serviced in (at least) three different types of sites:

1. run-through tracks, where simple processes may be executed;

2. service tracks, where locomotives are isolated from the main line, and more complex and lengthy processes (like repairs) may be undertaken;

3. main shops, where locomotives may be even disassembled.
The train delay due to the time spent refueling strongly depends on the site on which the fueling stop occurs. Fast refueling events (including on-road refueling operations performed by fueling trucks) are typically associated to run-through tracks while refueling events on service tracks and main shops require more time. Raviv and Kaspi [2012] assume that the train delay cost caused by refueling operations is equal to 2010US$ 250 (2008US$ 246.75) for each fueling event. This assumption is the same adopted in the problem solving competition “Locating Locomotive Refueling Stations” organized in 2010 by the Railway Applications Section of INFORMS, and won by Raviv and Kaspi. The cost of 2008US$ 246.75 assumed in Raviv and Kaspi [2012] is reasonable under the framework adopted in the competition:

1. the only source of fueling are the fueling trucks;
2. all the trains are pulled exactly by one locomotive;
3. assume instantaneous refueling time;
4. a train incurs a fixed cost if it is refueled.

According to Schafer and Barkan [2008] the industry expert opinion is that the cost of delay for a single train is in the range of 2006US$ 200 to 2006US$ 300 per hour (2008US$ 213.4 to 2008US$ 320.12 per hour). The fueling stop cost is determined by various characteristics of the train being delayed and of the yard. In particular, the characteristics of the train and the time spent refueling (and waiting to be refueled) determine the fueling stop cost. BNSF Railway Twin Cities division [2006] reports that refueling can take up to 10 hours in some congested rail yards (like the ones in Pasco-Washington, Seattle or Vancouver). For this reason BNSF (a U.S. railway company) has invested and continues to invest in new fueling facilities that are able to refuel a train in less than one hour (the station in Minot should be able to refuel a train in about 45 minutes). Souten et al. [2008] indicate an average refueling time of 1.5 hours for the BNSF yard located in San Bernardino. According to Gannett-Fleming [2008], estimates provided by Norfolk Southern (another U.S. railway company) indicated that the fueling operation typically takes 3 hours at the Dillerville Yard (Lancaster County, Pennsylvania). The uncertainty about the fueling stop cost is further increased by the different characteristics of trains, in particular by their priority. Given the same refueling time, high priority trains (being associated to a higher value of time) have a higher fueling stop cost with respect to low priority trains.

This study assumes that all the trains have the same priority and that the fueling stop cost is uniquely determined by the amount of time spent refueling (and waiting to be refueled). We also assume for the fueling
events a hourly cost equal to the one adopted for the idling events that is 2008US$ 111.51 according to Wilbur Smith Associates [2010]. According to GE Harris Energy Systems [2000], 80% of the locomotive fleet units are refueled on service tracks where each stop takes on average 6 hours, while the remaining 20% are refueled on run-through tracks, where each stop takes on average 30 minutes. Equivalently, we may assume that each fueling event takes on average 4.9 hours and has a fueling stop cost of 2008US$ 546.4. Figure 1 reports the number of fueling stops (left y axis) and the corresponding costs (right y axis) of 288 consist types (the numbers in the x axis) working 52 weeks, 50 hours per week.

![Figure 1: Yearly fueling events and cost for a single consist.](image)

### 5.3 The fueling homogeneity

Each consist type may present a certain grade of homogeneity according to one or more parameters used to characterize the locomotive types joined inside that consist type. Given the operative conditions, each locomotive type is characterized by its range or equivalently by the number of fueling stops required over a fixed time horizon (frequency of the fueling events). The frequency of the fueling events for a consist type is determined by the locomotive type with the shortest range (locomotives cannot run out of fuel). The locomotive types that form a consist type could have similar or dissimilar ranges. In a fueling heterogeneous consist, i.e. built using locomotives with dissimilar ranges, the locomotives with the longer ranges exploit
their fuel capacity inefficiently. In a fueling heterogeneous consist a portion of the fuel remains unused in the locomotives tanks. Thus, the money value of the fuel is not productively invested causing an opportunity cost that we denote as heterogeneity fueling cost \( (HFC) \). On the contrary, a group of locomotives types characterized by similar ranges has a low \( HFC \) and exploits the fuel tank capacity more efficiently. We may shortly define *perfectly homogeneous* a consist type characterized by a \( HFC = 0 \). The \( HFC \) of a consist is closer to zero the more *homogeneous* it is.

6 The heterogeneity fueling cost

Consider a locomotive type \( X \) (long range), joined with a locomotive type \( Y \) (short range) in a \( XY \) consist type. Since \( Y \) cannot run out of fuel, a portion of the fuel stored in the tank of \( X \) will not be exploited generating a heterogeneity fueling cost \( (HFC) \).

The \( HFC \) introduced in this study concurs in the consists selection process though active and ownership costs are dominant. Referring to 2008US$ we have:

1. the highest consist \( HFC \) is US$ 0.89 per day;
2. the lowest consist ownership cost is US$ 31.28 per hour;
3. the lowest consist active cost cost is US$ 80 per hour.

The \( HFC \) may become relevant only when we compare consist types with almost equivalent value of the sum (active cost + ownership cost).

6.1 Neglecting locomotive passive movements

According to Ahuja et al. [2005b] and Vaidyanathan et al. [2008a] the objective function of the LPP accounts for unused locomotives (savings) and costs due to:

1. active locomotives;
2. ownership;
3. deadheading and light-travelling locomotives (passive movements).
The consist type has a crucial role during the active part of a consist unit service while it is not relevant during passive movements (deadheading and light-traveling).

The objective function of the consists selection consider only fueling costs, active and ownership costs since the choice of the consist types is done looking at the motive performances requested during active movements (TE and HP constraints). Locomotive types do not matter during passive movements, for this reason passive movement costs are not considered in the consists selection. Table III (taken from John and Ahuja [2008]) confirms that for each locomotive type the active cost is significantly greater than the deadheading cost that, as expected, is the same for all the locomotive types (i.e. locomotive type is irrelevant during passive movements).

<table>
<thead>
<tr>
<th>Locomotive class</th>
<th>Active cost per hour (2008US$)</th>
<th>Deadheading cost per hour (2008US$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC4400CW</td>
<td>155</td>
<td>9</td>
</tr>
<tr>
<td>C40-8/C40-8W</td>
<td>125</td>
<td>9</td>
</tr>
<tr>
<td>SD40/SD40-2/SD40-3</td>
<td>105</td>
<td>9</td>
</tr>
<tr>
<td>ES44DC</td>
<td>125</td>
<td>9</td>
</tr>
<tr>
<td>GP40/GP40-2</td>
<td>80</td>
<td>9</td>
</tr>
</tbody>
</table>

Table III: Active and deadheading costs comparison.

### 6.2 Neglecting train to train connections

The train to train connections hold a crucial role in the LPP optimization but they are neglected in the consists selection. In a train to train connection, we use the same consist to serve several trains. In a sequence of trains, the train with the highest HP/Tonnage requirements determines the minimal performance of the shared consist. The performance of the chosen consist is suited for some trains of the sequence and excessive for others. In this study, we want to identify for each train the best suited consist type (while respecting the fleet size constraints), therefore the train to train connections are neglected.

Figure 2 (taken from Ahuja et al. [2005b]) shows the space-time network structure including ground nodes, ground arcs and connection arcs that are essential to model deadheading, high-traveling and train to train connection in the LPP, while Figure 3 shows the space-time network structure adopted in the consists selection model.
7 Models and Data

The consists selection model solves the weekly assignment of consists by neglecting passive movements (deadheading and light traveling) and train to train connections. To estimate the potential savings made by the adoption of the consists selection, we create realistic locomotives specification, trains specifications and schedules datasets and we apply the preliminary consists selection to realistic scenarios and instances derived from these datasets.

7.1 Mathematical modeling of the consists selection

We model the weekly consists selection as an integer multicommodity flow problem with side constraints on a spacetime network. Each consist type defines a commodity in this network. Neglecting locomotive passive movements and train to train connections we have a space-time network $G = (N, A)$ in which arcs denote trains and nodes denote events (departures and arrivals of trains).

The set of arcs $A$ coincides with the set of train arcs $TrArcs$. The set of nodes $N$ is formed by the two subsets:

1. arrival nodes ($ArrNodes$);
2. departure nodes ($DepNodes$).

Each train $l$ is characterized by the following attributes:
1. dep-time($l$), the departure time of a train $l$;

2. arr-time($l$), the arrival time of a train $l$;

3. dep-station($l$), the departure station of a train $l$;

4. arr-station($l$), the arrival station of a train $l$;

5. $T_l$, the tonnage requirement for a train $l$;

6. $HP_l$, the HP requirement for a train $l$.

Three different sets of locomotives may be associated with each train $l$:

1. most preferred[$l$], the preferred locomotive types;

2. less preferred[$l$], the accepted (paying a penalty) locomotive types;

3. prohibited[$l$], the not permitted locomotive types.

Given the set of all locomotive types $K$, $k$ denotes a particular locomotive type belonging to $K$.

Each locomotive type $k \in K$ is characterized by the following attributes:

1. $B^k$, the fleet-size of a locomotive of type $k$;

2. $h^k$, the horsepower (HP) of a locomotive of type $k$;

3. $t^k$, the tractive effort (TE) of a locomotive of type $k$;

4. $b^k$, the number of axles on a locomotive of type $k$;

5. $G^k$, the ownership cost of a locomotive of type $k$;

6. $c_{f}^{k}$, the cost of assigning an active locomotive of type $k$ to a train $l$.

The model relies on the following definitions:

1. $C$, the set of consist types available for assignments;

2. $c \in C$ denotes a specific consist type;
3. $\alpha_{ck}$, the number of locomotives of type $k \in K$ in a consist $c \in C$;

4. $I[i]$, the set of arcs entering in the node $i$;

5. $O[i]$, the set of arcs leaving the node $i$;

6. $c_f^l$, the cost of assigning an active consist of type $c \in C$ to a train arc $l$;

7. $e_f^l$, the heterogeneity fueling cost of a consist $c$ that pulls a train $l$ for its entire travel time;

8. $f_f^l$, the fueling stops cost of a consist $c$ that pulls a train $l$ for its entire travel time.

The decision variables are the following:

1. $z^c$, a binary variable which takes value 1 if a consist type $c \in C$ is used;

2. $x_f^l$, a binary variable which takes value 1 if a consist type $c \in C$ flows on arc $l \in TrArcs$.

Inspired by Ahuja et al. [2005b] and Vaidyanathan et al. [2008a], we solve a weekly consist assignment with a fixed number $p$ of maximum available consist types. The value of $p$ should be low to have LPP solutions manageable and useful in real-life applications.

\[
\min \ : w = \sum_{l \in TrArcs} \sum_{c \in C} c_f^l x_f^l + \sum_{l \in TrArcs} \sum_{c \in C} \sum_{k \in K} \alpha_{ck}^l G_k^l + \sum_{l \in TrArcs} \sum_{c \in C} [e_f^l x_f^l + f_f^l x_f^l]
\]

subject to

\[
\sum_{c \in C} \sum_{k \in K} \alpha_{ck}^l x_f^l \geq T_l, \quad \forall l \in TrArcs
\]

\[
\sum_{c \in C} \sum_{k \in K} \alpha_{ck}^l x_f^l \geq HP_l, \quad \forall l \in TrArcs
\]

\[
\sum_{c \in C} x_f^l = 1, \quad \forall l \in TrArcs
\]

\[
\sum_{l \in TrArcs} \sum_{c \in C} \alpha_{ck}^l x_f^l \leq B_k, \quad \forall k \in K
\]

\[
z^c \geq x_f^l, \quad \forall l \in TrArcs, \ c \in C
\]

\[
\sum_{c \in C} z^c \leq p, \quad p = 3, 5, 7, \ldots, 17
\]

\[
x_f^l \in \{0, 1\}, \quad \forall l \in TrArcs, \ c \in C
\]

\[
z^c \in \{0, 1\}, \quad \forall c \in C
\]

\[
\sum_{c \in C} x_f^l b^c \leq 24, \quad \forall l \in TrArcs
\]
The structure of our model has some similarities with the ones of the models proposed by Ahuja et al. [2005b] and Vaidyanathan et al. [2008a] but presents some important differences. The models proposed in Ahuja et al. [2005b] and Vaidyanathan et al. [2008a] are developed to solve the daily consists assignment problem, a first step toward the solution of the very large scale weekly assignment problem. Our model is able to solve directly the weekly problem because the focus is on the preliminary consists selection, a problem where some extremely complex aspects of the LPP (passive movements, train to train connections) are irrelevant. The objective function takes into account the active cost (as done in Ahuja et al. [2005b] and Vaidyanathan et al. [2008a]) and the ownership cost (as done explicitly in Ahuja et al. [2005b] and implicitly in Vaidyanathan et al. [2008a]). Our model introduces in the objective function the fueling costs term (i.e. \[ \sum_{l} \sum_{c} [e_{c}^{l}x_{c}^{l} + f_{c}^{l}x_{c}^{l}] \]) that were not considered in Ahuja et al. [2005b] and Vaidyanathan et al. [2008a].

The models have in common the following constraints:

1. constraint (1b) ensures that for each train \( l \) the tractive effort requirement is satisfied;
2. constraint (1c) ensures that for each train \( l \) the horsepower requirement is satisfied;
3. constraint (1d) ensures that to every train \( l \) is assigned exactly one active consist;
4. constraint (1e) ensures that the number of used locomotives units of each locomotive type is not greater than the available one (fleet size constraint).

The constraint (1f) counts the number of used consist types while the constraint (1g) limits (up to \( p \)) the maximum number of consist types available in the solution of the consists selection problem. The last constraint (1j) impacts the set of accepted consist types imposing a maximum number of 24 active axles per consist. In this study we consider 7 locomotive types that generate the general set of accepted locomotive combinations (consist types) from which we extract the set \( C \). This last constraint may be implicitly handled excluding from the general set of accepted consist types the consist types that have more than 24 active axles.

### 7.2 Assessment of savings achievable introducing the consists selection

In the previous studies the LPP optimization was performed taking the set \( C \) of available consist types as given. Our objective is to assess the savings achievable introducing the consists selection before the LPP optimization. Given an instance of the LPP, the exact valuation of the achievable savings may be obtained after solving the LPP (and determining the train to train connections) and finally updating the LPP solution.
to honor the fueling constraints in the routing phase. In fact, the exact valuation of savings achievable introducing the consists selection may be obtained only after having solved the LPP and the LRP (in which fueling constraints are honored). Ideally we should integrate the consists selection in the solution procedures developed in Vaidyanathan et al. [2008a] and Vaidyanathan et al. [2008b]. Another important aspect is that, according to Nahapetyan et al. [2007], actual train and locomotive schedules suffer from disruptions and delays (30÷40 of trains does not depart with a set of locomotives specified in the LPP solution and the percentage of trains that arrive ontime is around 50%). Thus, it is quite difficult to evaluate the actual savings achievable introducing the consists selection even knowing the LPP solution.

Due to these complex aspects and to the lack of essential information, this study propose a different approach to estimate the potential savings achievable through the consists selection. To summarize, our objective is to estimate the potential savings achievable adopting a consist types set \( C \) obtained accounting for the active cost, the ownership cost, the HFC and the FSC with respect to a consist types set \( C \) obtained accounting only for active and ownership costs. To achieve this objective we solve a reference model in which the consists selection is performed considering only the active cost and the ownership cost. The reference model is named model MR and the consists selection model is named model MS.

In the reference model (model MR) the objective function accounts only for active and ownership costs:

\[
min: w = \sum_{l \in TrArcs} \sum_{c \in C} c_l^c x_l^c + \sum_{l \in TrArcs} \sum_{c \in C} \sum_{k \in K} \alpha_{ck}^c x_l^c G_k
\]  

(2a)

In the consists selection model (model MS) the objective function accounts also for the HFC and FSC terms (i.e. \( \sum_{l} \sum_{c} [e_{c}^l x_{l}^c + f_{c}^l x_{l}^c] \)). The assessment of savings is completed solving these two models and comparing the two respective overall cost (active + ownership + HFC + FSC). The constraints are the same for the two models.

The objective value of the model MS represents an overall cost (since the HFC and the FSC are present in the objective function). To obtain the overall cost of the model MR we add the objective value of MR (i.e. the active + ownership costs) with the HFC and the FSC that correspond to the MR solution.

The difference between this MR overall cost and the objective value of MS provides an estimate of the weekly savings obtained accounting for the HFC and the FSC in the optimization program.
The model MS may have several optimal equivalent solutions, all characterized by the same overall cost (that coincides with the MS objective value). On the contrary, MR may have several optimal solutions that are all equivalent in terms of optimal MR objective value but may present different \([HFC + FSC]\) values (and so different overall costs) since the \(HFC\) and the \(FSC\) are not accounted in MR optimization (they are calculated ex-post once we know the MR solution).

The assessment of savings is obtained through a three step procedure. In the first step we solve the models MR and MS: the optimal MS solutions are characterized by the minimum achievable overall cost (active + ownership + \(HFC + FSC\)). For each instance, in general, the optimal MR solution is not unique and we found a set of optimal equivalent solutions characterized by the same active + ownership costs (same optimal objective value) but different \(HFC\) and \(FSC\) (not considered in the MR objective function).

In the second step, for each instance, we identify the optimal MR solutions characterized by the highest \([HFC + FSC]\). Then, for each instance, we calculate the overall cost as the sum of the optimal MR objective value and the highest value of \([HFC + FSC]\) (we refer to this sum as the MR overall cost).

In the third step, for each instance, we calculate the difference between the MR overall cost and the MS overall cost obtaining an estimate of the maximum weekly savings achievable implementing the consists selection.

### 7.3 The dataset

Vaidyanathan et al. [2008a] solve the LPP implementing the consist flow formulation in two different scenarios (with similar size) provided by CSX:

\(<388\text{ trains, 6 locomotive types}\>\text{ and } <382\text{ trains, 6 locomotive types}\>

For each scenario the LAP is solved finding the total number of locomotives used in 8 sub-scenarios identified by 8 different consist types sets \(C\) (that contain from 3 to 17 consist types).

Railway companies do not provide such kind of detailed data without a partnership. Nevertheless, a deep search on scientific publications, economic and technical reports, manuals and other freely available sources, give us the realistic data needed for our analysis. The information retrieved may be grouped in four categories:
1. locomotives and rolling stock data (train cars data);

2. train data;

3. consist data;

4. tracks data.

To obtain a realistic set of consist types $C$, the proportion of train types in the set of train services matter more than the number of train services: a set of 1000 grain trains (a flat, dull train services set) would produce an unrealistic composition of $C$. Thereby, we create a set of 229 train services characterized by a realistic proportion of the following 3 different train speed classes:

1. auto trains (10 trains);

2. intermodal trains (65 trains);

3. merchandize and bulk trains (154 trains).

The 229 train services are created to represent a realistic set of train for a typical East-coast railway company (it may be CSX for example). On average, a West-coast company would have longer routed distances, higher travel times, higher weights of the trains and a different distribution of trains among the three speed classes auto, intermodal and merchandize.

Table IV lists the information sources exploited.
<table>
<thead>
<tr>
<th>Locomotives and train cars data</th>
<th>Train data</th>
<th>Consist data</th>
<th>Tracks data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parajuli [2005]</td>
<td>Lai et al. [2008]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammah-Tago [2006]</td>
<td>Schafer and Barkan [2008]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSX Corporation [2006]</td>
<td>Souten et al. [2008]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hawthorne et al. [2006]</td>
<td>Brosseau and Ede [2009]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireson [2007]</td>
<td>Innovative Scheduling [2009]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lindsey [2007]</td>
<td>Roucolle and Elliott [2010]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>John and Ahuja [2008]</td>
<td>Raviv and Kaspi [2012]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaidyanathan et al. [2008a]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brosseau and Ede [2009]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSX Corporation [2009]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICF International [2009]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Innovative Scheduling [2009]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metrolinx [2010]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilbur Smith Associates [2010]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV: Data sources.
Our instances are generated considering two sets of locomotive types, the first one is the same adopted by Vaidyanathan et al. [2008a] and includes 6 locomotive types (AC4400CW, AC6000CW, C40-8, GP40-2, SD40-2, SD60I), the second one is obtained adding the locomotive type ES44DC to the previous group of 6 locomotive types.

Each locomotive type may be a preferred, accepted or prohibited choice for the 3 different train speed classes and Innovative Scheduling [2009] offers a realistic reference on this subject. Knowing the preferred, accepted and prohibited \(\langle\text{train class, locomotive type}\rangle\) connections, we may build the set of allowed consist types for each train speed class. Combining up to 7 locomotive types, we obtain 288 valid consist types and the corresponding prohibited connections:

1. 36 consist types are prohibited for merchandise and bulk trains;
2. 218 consist types are prohibited for auto trains;
3. 249 consist types are prohibited for intermodal trains.

The prohibited \(\langle\text{train class, locomotive type}\rangle\) connections reduce the set of accepted locomotive combinations (valid consist types). For example, the locomotive type F is prohibited for intermodal and auto trains and is allowed for merchandise trains while the locomotive type A is allowed for intermodal and auto trains and is prohibited for merchandise trains. Therefore, a consist type obtained joining locomotive types A and F is of no use. Considering all these restrictions, it is possible to obtain a set of 288 valid consist types to be analyzed in the consists selection phase.

The data extracted from the selected information sources have been integrated in a simulation program used to generate our instances. Table V shows some relevant CSX locomotive data.

In particular, Table V reports (in the last three columns) the active costs multiplication factors of the preferred and accepted connections between locomotive types and train speed classes. The basic active costs (reported in the fourth column) multiplied by these coefficients provides the actual active cost per hour (in 2008US$). To facilitate the description of the consist types we adopt an alphabetic code for each locomotive type (as reported in the second column).

We extract 32 different instances from the dataset used to generate the train schedules. These instances are grouped in four different scenarios (we have 8 instances in each scenario), each scenario is characterized by three different parameters:
<table>
<thead>
<tr>
<th>Locomotive type</th>
<th>Locomotive type alpha code</th>
<th>HP (US $ per hour)</th>
<th>Active cost$ (US $ per hour)</th>
<th>Lease cost$ (US $ per hour)</th>
<th>Ownership cost$ (US $ per hour)</th>
<th>Intermodal trains(^c)</th>
<th>Auto trains(^c)</th>
<th>Merchandize, Bulk trains(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC4400CW</td>
<td>A</td>
<td>4400</td>
<td>155</td>
<td>28</td>
<td>43.792</td>
<td>1</td>
<td>1.2</td>
<td>Prohibited</td>
</tr>
<tr>
<td>AC6000CW</td>
<td>B</td>
<td>6000</td>
<td>155</td>
<td>28</td>
<td>43.792</td>
<td>1</td>
<td>1.2</td>
<td>Prohibited</td>
</tr>
<tr>
<td>C40-8</td>
<td>C</td>
<td>4000</td>
<td>125</td>
<td>26</td>
<td>40.664</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>ES44DC</td>
<td>D</td>
<td>4400</td>
<td>125</td>
<td>29</td>
<td>45.356</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>GP40-2</td>
<td>E</td>
<td>3000</td>
<td>80</td>
<td>20</td>
<td>31.28</td>
<td>Prohibited</td>
<td>Prohibited</td>
<td>1</td>
</tr>
<tr>
<td>SD40-2</td>
<td>F</td>
<td>3000</td>
<td>105</td>
<td>20</td>
<td>31.28</td>
<td>Prohibited</td>
<td>Prohibited</td>
<td>1</td>
</tr>
<tr>
<td>SD60I</td>
<td>G</td>
<td>3800</td>
<td>117</td>
<td>24</td>
<td>37.536</td>
<td>Prohibited</td>
<td>1.2</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Table V: Locomotive types characteristic data.

\(^a\) 2008USS, John and Ahuja [2008], figures for AC6000CW and SD60I are a guesswork
\(^b\) estimated values, derived from the lease cost according to Wilbur Smith Associates [2010]
\(^c\) the additional 20% of penalty cost is applied if an accepted connection is used instead of a preferred one, Liu [2003]

1. size of the locomotives fleet (actual fleet size or reduced fleet size).

2. single locomotive consists (allowed or prohibited);

3. number of available locomotive types (6 or 7);

The actual locomotives fleet size is the one of the 2011 CSX locomotives fleet (Table VI).

<table>
<thead>
<tr>
<th>Locomotive class</th>
<th>Units 2005(^a)</th>
<th>Units 2006(^b)</th>
<th>Units 2011(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC4400CW</td>
<td>593</td>
<td>593</td>
<td>621</td>
</tr>
<tr>
<td>C40-8/C40-8W</td>
<td>532</td>
<td>532</td>
<td>529</td>
</tr>
<tr>
<td>SD40/SD40-2/SD40-3</td>
<td>404</td>
<td>402</td>
<td>529</td>
</tr>
<tr>
<td>AC6000CW</td>
<td>116</td>
<td>117</td>
<td>117</td>
</tr>
<tr>
<td>ES44DC</td>
<td>0</td>
<td>100</td>
<td>302</td>
</tr>
<tr>
<td>SD60I/SD60/SD60M</td>
<td>90</td>
<td>90</td>
<td>94</td>
</tr>
<tr>
<td>GP40/GP40-2</td>
<td>0</td>
<td>0</td>
<td>416</td>
</tr>
</tbody>
</table>

Table VI: The 7 locomotive types in the 2005, 2006 and 2011 CSX fleets.

\(^a\) CSX Corporation [2005]  
\(^b\) CSX Corporation [2006]  
\(^c\) 2011 data source: www.thedieselshop.us/CSX.HTML (accessed April 2011)
The reduced fleet size is obtained considering 25% of the actual size (for each one of the 7 locomotive type groups the reduced sizes are rounded suitably).

Due to the risk of track block, CSX strongly discourage the assignment of consists composed by only one locomotive (single locomotive consist), and a penalty for this kind of assignment is adopted in Ahuja et al. [2005b]. In the solution obtained by Vaidyanathan et al. [2008a] the single locomotive consists are not adopted. In this study we consider two dichotomous alternatives: single locomotive consists are allowed without any penalty or are not allowed at all.

The four scenarios considered in this study are the following:

1. 6 locomotive types; single locomotive consists prohibited; actual fleet size;
2. 6 locomotive types; single locomotive consists prohibited; reduced fleet size;
3. 7 locomotive types; single locomotive consists allowed; actual fleet size;
4. 7 locomotive types; single locomotive consists allowed; reduced fleet size.

8 Numerical results and Discussion

This section shows the potential savings achievable by implementing the preliminary consists selection. We report the savings in terms of money (2008US$), number of fueling stops and servicing hours saved. Unless otherwise stated, we have expressed the monetary values in terms of 2008US$ throughout all the current paper.

8.1 Results

The results obtained by us rely on some simplifying assumptions. The locomotive types ranges are calculated relying on the locomotive utilization profile (the breakdown of locomotives activity within a 24-hour period). More precisely, we rely on the partition of the engine operative service time, i.e. the percentage of time that the diesel engine is turned on and consumes fuel, within a representative (based on yearly averages) 24-hour period.

We focus on the locomotive duty cycle i.e. the profile of the different locomotive power settings (Idle, Notch
levels 1 through 8) as percentages of the engine operating time. To figure out the FSC, we assume that each locomotive type has the same duty cycle of a representative mainline freight locomotive reported in Railway Association of Canada [2008]. This representative duty cycle is determined by evaluating the time spent at each power notch level for a statistically significant sample of locomotives. In other words, we assume that all the locomotive types spend their operative service time at each engine power level (notch level) in the same manner of the representative mainline freight locomotive. Moreover, for each locomotive type, we associate each notch level with the corresponding fuel consumption rate that is specific for each locomotive type (Seedah and Harrison [2010], ENVIRON International Corporation [2007]).

The yearly opportunity cost due to the unexploited fuel is calculated as the yearly total return that a railways company would obtain investing the value of the fuel (immobilized in the tanks of the consists) at the beginning of the year. Namely, we assume an annual real total return equal to 6.5%, this value is in line with the average of the Barclays Capital U.S. Aggregate Bond Index in the period 1982-2008 (9.45% according to Barclays Capital [2011]) discounted by an average inflation of about 3% (see for example http://www.multpl.com/inflation/table).

Figure 4 shows that the oil price and consequently the diesel fuel price were exceptionally high in 2008.

Figure 4: Average U.S. crude oil and diesel retail prices in the period 1980 - 2012.
To avoid opportunity cost overvaluation, we adopt the average price of diesel fuel in the period January 2008 — August 2012 (2008US$ 2.68 according to CSX Corporation [2011a], CSX Corporation [2011d], CSX Corporation [2011b], CSX Corporation [2011c] and http://www.eia.gov/petroleum/gasdiesel/).

Having calculated the $FSC$ and the $HFC$ for each consist type, we implement the consists selection solving the two models MR and MS in the four scenarios with 8 instances for each scenario (associated with the 8 values of $p \leq 3, 5, \ldots, 17$ in the constraint (1g)). To identify the 64 possible solutions (32 for each model MR and MS) we adopt a compact notation, to give an example MRNo_25%-03 indicates the consists selection solution obtained applying the model MR to the scenario with 6 locomotive types and single locomotive consist prohibited (No to the ES44DC type and No single locomotive consist), with the reduced fleet size (25% of the actual fleet) and with $p \leq 3$. Analogously, MSYes_100%-17 means model MS, 7 locomotives types (type ES44DC permitted) and Yes to single locomotive consist, actual fleet size available (100% of the fleet) and $p \leq 17$. The models have been solved with CPLEX 12.2 on a Core 2 Quad Q9550 2.83 Ghz and 4 Gb RAM. Figures 5 and 7 report the maximum achievable yearly savings (in US2008$) while Figures 6 and 8 show the number of locomotives and consists required in the Yes and No scenarios respectively. In Figures 6 and 8, the left y axis refers to the number of used consist types, while the right one refers to the number of used locomotive units (that form the consist units).

![Graph](image.png)

Figure 5: Yearly savings MR VS MS, Yes single locomotives & Yes ES44DC type.
Figure 6: # Consists and Locomotives used MR VS MS, Yes single locomotives & Yes ES44DC type.

Figure 7: Yearly savings MR VS MS, No single locomotives & No ES44DC type.
Figure 8: # Consists and Locomotives used MR VS MS, No single locomotives & No ES44DC type.

Considering 100% of the actual fleet size, for the instances characterized by 7 locomotive types and single locomotive consists allowed (MRYes, MSYes) we may observe the following:

1. MR and MS produce essentially the same requirement in terms of needed fleet size; the only difference is between MRYes17-100% and MSYes17-100% (382 and 381 locomotives respectively);

2. all the considered instances have a feasible solution;

3. the number of used locomotives remains stable ranging from 380 for \( p \leq 17 \) to 389 for \( p \leq 3 \);

4. the number of used consist types coincides with its maximum permitted value for both MR and MS except for MR when \( p \leq 17 \), in this case it is equal to 15;

5. the maximum achieved yearly savings are approximately 2008US$ 110000.

In the same scenario (MRYes, MSYes), if we consider 25% of the actual fleet size we observe the following:

1. MR and MS produce the same requirement in terms of needed fleet size;

2. the instances with \( p \leq 3 \) are infeasible;
3. the number of used locomotive varies ranging from 388 for \( p \leq 17 \) to 475 for \( p \leq 5 \);

4. the number of used consist types coincides with its maximum permitted value \( p \) and is the same for MR and MS;

5. the maximum achieved yearly savings are approximately 2008US$ 92000.

We do not observe any effect of the consists selection on the fleet size in the 25% scenario and very small effect in the 100% scenario (a small difference between MRYes17-100% and MSYes17-100% i.e. 382 and 381 used locomotives respectively). Thus, as expected, the introduction of the fueling cost terms \( FSC \) and \( HFC \) in the LPP optimization essentially does not impact the number of used locomotives.

In the ”25% scenario”, for both models MR and MS, the number of used consist types coincides with the maximum permitted value also when \( p \leq 17 \): the reduced availability of the best choices imposes the utilization of second choices (exploiting all the 17 available consist types). In the same way, we summarize the results for the instances characterized by 6 locomotive types and single locomotive consists prohibited (MRNo, MSNo). If we consider 100% of the actual fleet size we observe the following:

1. MR and MS produce the same results in terms of needed fleet size;

2. the number of used locomotives decreases passing from 568 to 499 when \( p \) increases;

3. the number of used consist types does not coincide with its maximum permitted value when \( p \leq 11,13,15,17 \), being equal to 10 and 11 for MR and MS respectively;

4. the maximum achieved yearly savings are approximately 2008US$ 108000.

If we consider 25% of the actual fleet size we observe that:

1. MR and MS produce the same results in terms of needed fleet size.

2. the instances with \( p \leq 3,5 \) are infeasible;

3. the number of used locomotives remains almost constant passing from 509 to 507 when \( p \) increases;

4. the number of used consist types coincides with its maximum permitted value for both MR and MS except when \( p \leq 15,17 \); when \( p \leq 15 \) it is equal to 15 for MR and 14 for MS, when \( p \leq 17 \) it is equal to 16 for MR and 14 for MS;
5. the maximum achieved yearly savings are approximately 2008US$ 9529.

Again, we do not observe effects of the consists selection on the needed fleet size. This is exactly what we expect because active and ownership costs (that depend on the number of used locomotive units) dominate $FSC$ and $HFC$. Comparing the Yes and the No scenarios, we observe that in the No instances the absence of single locomotive consists reduces the LAP optimization possibilities and the solution flexibility leading to:

1. an increased average consist size ⇒ increased number of used locomotives;
2. an increased consist size constraints tightness ⇒ more infeasible solutions;
3. a reduced consist types availability ⇒ reduced number of (useful and) used consist types;
4. smaller savings when fleet size constraints are tight (25% scenario);
5. consistently bigger savings when fleet size constraints are loose (100% scenario).

The differences are even greater comparing 100% and 25% scenarios. We observe that in 25% instances the strongly reduced availability of locomotives reduces the optimization possibilities:

1. the best consist types are used up rapidly ⇒ utilization of second choices that increases the number of used consist types;
2. utilization of second choices (costly consist types) ⇒ higher average size of consists i.e. more used locomotives;
3. significantly smaller savings.

Figures 9, 10, 11, and 12 report the distribution of savings among the three different train speed classes (auto, intermodal, merchandize) in the four scenarios Yes100%, Yes25%, No100%, and No25%. Comparing the yearly savings distribution that characterize Yes and No instances we observe several similarities between Yes100% and No100% histograms while these similarities disappear in the Yes25% and No%25 histograms. The differences in the yearly savings distributions are even more evident among the 100% and 25% instances. This fact evidences that the tight size constraints impact the solution (and the savings opportunities) more than the unavailability of the sigle locomotive consists and of the (additional) consist type ES44DC.
Figure 9: Yearly savings histograms for the Auto, Intermodal and Merchandise trains in the Yes100% instances.

Figure 10: Yearly savings histograms for the Auto, Intermodal and Merchandise trains in the Yes25% instances.
Figure 11: Yearly savings histograms for the Auto, Intermodal and Merchandize trains in the No100% instances.

Figure 12: Yearly savings histograms for the Auto, Intermodal and Merchandize trains in the No25% instances.
Table VII evidences two groups of consist type changes: profitable and unprofitable (savings < 0).

<table>
<thead>
<tr>
<th>Changes</th>
<th>All Count</th>
<th>All Percent</th>
<th>100% Count</th>
<th>100% Percent</th>
<th>25% Count</th>
<th>25% Percent</th>
<th>Yes Count</th>
<th>Yes Percent</th>
<th>No Count</th>
<th>No Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC→AC</td>
<td>165</td>
<td>42.42%</td>
<td>163</td>
<td>65.46%</td>
<td>2</td>
<td>1.43%</td>
<td>70</td>
<td>30.70%</td>
<td>95</td>
<td>59.01%</td>
</tr>
<tr>
<td>BCC→ACC</td>
<td>84</td>
<td>21.59%</td>
<td>84</td>
<td>33.73%</td>
<td>0</td>
<td>0.00%</td>
<td>28</td>
<td>12.28%</td>
<td>56</td>
<td>34.78%</td>
</tr>
<tr>
<td>EFF→EEF</td>
<td>47</td>
<td>12.08%</td>
<td>0</td>
<td>0.00%</td>
<td>47</td>
<td>33.57%</td>
<td>47</td>
<td>20.61%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>D→C</td>
<td>8</td>
<td>2.06%</td>
<td>0</td>
<td>0.00%</td>
<td>8</td>
<td>5.71%</td>
<td>8</td>
<td>3.51%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>CF→FF</td>
<td>3</td>
<td>0.77%</td>
<td>0</td>
<td>0.00%</td>
<td>3</td>
<td>2.14%</td>
<td>0</td>
<td>0.00%</td>
<td>3</td>
<td>1.86%</td>
</tr>
<tr>
<td>D→F</td>
<td>2</td>
<td>0.51%</td>
<td>0</td>
<td>0.00%</td>
<td>2</td>
<td>1.43%</td>
<td>2</td>
<td>0.88%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>AC→CC</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>1</td>
<td>0.44%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>AA→AC</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>1</td>
<td>0.44%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>C→G</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>1</td>
<td>0.44%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>CC→GG</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.62%</td>
</tr>
<tr>
<td>EF→EE</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.62%</td>
</tr>
<tr>
<td>Count</td>
<td>314</td>
<td>247</td>
<td>67</td>
<td>158</td>
<td>161</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Changes</th>
<th>All Count</th>
<th>All Percent</th>
<th>100% Count</th>
<th>100% Percent</th>
<th>25% Count</th>
<th>25% Percent</th>
<th>Yes Count</th>
<th>Yes Percent</th>
<th>No Count</th>
<th>No Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>E→F</td>
<td>57</td>
<td>14.65%</td>
<td>0</td>
<td>0.00%</td>
<td>57</td>
<td>40.71%</td>
<td>57</td>
<td>25.00%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>CE→DE</td>
<td>8</td>
<td>2.06%</td>
<td>0</td>
<td>0.00%</td>
<td>8</td>
<td>5.71%</td>
<td>8</td>
<td>3.51%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>CF→CC</td>
<td>3</td>
<td>0.77%</td>
<td>0</td>
<td>0.00%</td>
<td>3</td>
<td>2.14%</td>
<td>0</td>
<td>0.00%</td>
<td>3</td>
<td>1.86%</td>
</tr>
<tr>
<td>AC→AA</td>
<td>2</td>
<td>0.51%</td>
<td>0</td>
<td>0.00%</td>
<td>2</td>
<td>1.43%</td>
<td>2</td>
<td>0.88%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>EEE→CD</td>
<td>1</td>
<td>0.26%</td>
<td>1</td>
<td>0.40%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.44%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>CG→CC</td>
<td>1</td>
<td>0.26%</td>
<td>1</td>
<td>0.40%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.44%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>EF→FF</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.62%</td>
</tr>
<tr>
<td>G→C</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>1</td>
<td>0.44%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>GG→CC</td>
<td>1</td>
<td>0.26%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.71%</td>
<td>0</td>
<td>0.00%</td>
<td>1</td>
<td>0.62%</td>
</tr>
<tr>
<td>Count</td>
<td>75</td>
<td>2</td>
<td>73</td>
<td>70</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>389</td>
<td>249</td>
<td>140</td>
<td>228</td>
<td>161</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VII: Consist type changes associated with savings and losses (savings < 0).
The presence of consist type changes that produce losses is justified by the scarcity of specific consist types: without fleet size constraints, i.e. unlimited locomotives availability, the optimization program would not consider this disadvantageous changes. In fact, these changes are exploited to free up some specific locomotive types to be used in the creation of profitable consist changes (otherwise impossible) that off set the losses of the disadvantageous changes and produce higher savings.

We note that on a total number of 389 consist changes we have only one occurrence of a consist change that involves consists with a different size (EEE→CD): the consists selection optimization program does not consider consists with equivalent performances but different sizes as substitutes due to the relevant active and ownership costs for each single locomotive.

The differences among the scenarios are particularly marked between 100% and 25% scenarios, specially for the distribution of consist changes that produce losses. For example, on a total number of 75 consist changes that cause a monetary loss, 73 of these changes are distributed among the instances of the 25% scenario. Figure 13 shows an example of savings and losses of the two groups of consist changes along with the travel time (to enhance the readability of the chart we have excluded the four less numerous consist changes in each group).

Figure 13: An example of savings and losses achievable with different consist changes in different travel times.
We conclude the results section with the Table VIII that reports the savings in terms of number of fueling stops and servicing hours.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No_25%-03le</td>
<td>infeasible</td>
<td>infeasible</td>
<td>infeasible</td>
<td>infeasible</td>
</tr>
<tr>
<td>No_25%-05le</td>
<td>infeasible</td>
<td>infeasible</td>
<td>infeasible</td>
<td>infeasible</td>
</tr>
<tr>
<td>No_25%-07le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No_25%-09le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No_25%-11le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No_25%-13le</td>
<td>4.89</td>
<td>254.26</td>
<td>2.28</td>
<td>0.47</td>
</tr>
<tr>
<td>No_25%-15le</td>
<td>119.05</td>
<td>6190.61</td>
<td>55.52</td>
<td>11.33</td>
</tr>
<tr>
<td>No_25%-17le</td>
<td>183.24</td>
<td>9528.62</td>
<td>85.45</td>
<td>17.44</td>
</tr>
<tr>
<td>No_100%-03le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No_100%-05le</td>
<td>1728.65</td>
<td>89889.71</td>
<td>806.11</td>
<td>164.51</td>
</tr>
<tr>
<td>No_100%-07le</td>
<td>1093.84</td>
<td>56879.93</td>
<td>510.09</td>
<td>104.1</td>
</tr>
<tr>
<td>No_100%-09le</td>
<td>1093.84</td>
<td>56879.93</td>
<td>510.09</td>
<td>104.1</td>
</tr>
<tr>
<td>No_100%-11le</td>
<td>2068.51</td>
<td>107562.58</td>
<td>964.6</td>
<td>196.86</td>
</tr>
<tr>
<td>No_100%-13le</td>
<td>2068.51</td>
<td>107562.58</td>
<td>964.6</td>
<td>196.86</td>
</tr>
<tr>
<td>No_100%-15le</td>
<td>2068.51</td>
<td>107562.58</td>
<td>964.6</td>
<td>196.86</td>
</tr>
<tr>
<td>No_100%-17le</td>
<td>2068.51</td>
<td>107562.58</td>
<td>964.6</td>
<td>196.86</td>
</tr>
<tr>
<td>Yes_25%-03le</td>
<td>infeasible</td>
<td>infeasible</td>
<td>infeasible</td>
<td>infeasible</td>
</tr>
<tr>
<td>Yes_25%-05le</td>
<td>1765.43</td>
<td>91802.3</td>
<td>823.27</td>
<td>168.01</td>
</tr>
<tr>
<td>Yes_25%-07le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_25%-09le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_25%-11le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_25%-13le</td>
<td>283.87</td>
<td>14761.4</td>
<td>134.55</td>
<td>27.46</td>
</tr>
<tr>
<td>Yes_25%-15le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_25%-17le</td>
<td>29.61</td>
<td>1539.75</td>
<td>14.03</td>
<td>2.86</td>
</tr>
<tr>
<td>Yes_100%-03le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_100%-05le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_100%-07le</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Yes_100%-09le</td>
<td>1728.17</td>
<td>89864.63</td>
<td>805.89</td>
<td>164.47</td>
</tr>
<tr>
<td>Yes_100%-11le</td>
<td>1093.84</td>
<td>56879.93</td>
<td>510.09</td>
<td>104.1</td>
</tr>
<tr>
<td>Yes_100%-13le</td>
<td>1093.84</td>
<td>56879.93</td>
<td>510.09</td>
<td>104.1</td>
</tr>
<tr>
<td>Yes_100%-15le</td>
<td>1370.31</td>
<td>71255.98</td>
<td>639.01</td>
<td>130.41</td>
</tr>
<tr>
<td>Yes_100%-17le</td>
<td>2113.04</td>
<td>109877.93</td>
<td>985.36</td>
<td>201.09</td>
</tr>
</tbody>
</table>

Table VIII: MR vs MS yearly savings in 2008US$, fueling hours and fuel stops.
9 Discussion

Figure 14 confirms the conjecture: the longer the travel, the higher the number of fueling events, the higher the achievable savings. Another aspect emerges analyzing the results: when we consider a suitable consists fleet (100% scenario), the savings concentrate on a small number of consist changes (BC→AC and BCC→ACC) and on a specific train service class (intermodal trains). The changes BC→AC and BCC→ACC are the most present because consist types AC and BC are highly interchangeable. Their active and ownership costs are equal and the performance differences between the locomotives AC4400CW and AC6000CW (A and B respectively) are small enough to keep satisfied (for a large part of the train set) the train service requirements. It is less evident why these savings concentrate on intermodal trains. Each consist type has different ownership and active hourly costs. Considering the preferred consist type for each train type (preferred connection) we may identify the minimum possible active cost that characterize each consist type. Summing this active cost with the ownership cost we obtain the minimum active and ownership cost per hour (ActOwn minimum cost per hour). Figure 14 shows how the performance of a consist in terms of TE and HP vary along with the ActOwn minimum cost.

Figure 14: TE and HP versus (Actual + Ownership) hourly cost.

The TE diminishes slowly and remains quite stable while the HP diminishes more rapidly when the
minimum cost decreases. Thus, if a wide range of HP performance is acceptable (as for slow trains), we could reduce the ActOwn cost preserving very similar performance in terms on TE. In this case several consist types with different ActOwn costs may be considered perfect substitutes, and the differences in terms of ActOwn cost may be very high and dominate the (much smaller) savings achieved by a long range homogeneous consists fleet. However, if the HP performance is critical (as for fast trains, like inter-modal that are the faster ones) we expect that only consist with similar HP (and so similar ActOwn cost) are perfectly interchangeable. Thereby, differences between interchangeable consist in terms of ActOwn are small, and savings offered by long range homogeneous consists become significant. A fleet of consist may offer significant savings just replacing the locomotive type B with the A (whenever possible). Furthermore since savings concentrate on (long travel) intermodal trains these changes are even more easy to implement because they involve only a portion of the consist fleet (the part that serves these specific trains) reducing the cost of the fleet renovation.

10 Conclusions and future work

In this study we propose a methodological innovation that is able to partially integrate LAP planning and routing phases. Our objective is to obtain LPP solutions that make the routing phase easier to handle and more economical. We pursue this objective considering the LPP in its consist flow formulation and accounting for information about consist characteristics, such as ranges and efficiency in the exploitation of fuel capacity, not featured in the previous studies. We focus on the identification of the set of initially available consist types to be used in the LAP optimization. This set were typically assumed as settled in terms of consist types (i.e. in terms of its qualitative composition) by the expertise of locomotive managers and the LPP were solved to identify the quantitative composition of the set. In this study, we propose an optimization model (consists selection) to identify the qualitative composition of the set of consist types to be used in the LPP optimization. We introduce the concept of consist fueling homogeneity, and we implement the preliminary consists selection that precedes the LPP optimization. This phase could identify consist types that are not captured by a purely cost-oriented consists selection but that may reduce the opportunity costs linked with the unexploited portion of the fuel stocks and simplify the consists fueling routing reducing the number of fueling stops (and the corresponding costs).

We consider several realistic instances, and we obtain yearly savings up to 201 fueling events (985 servicing
hours, 2008US$ 110000). We found that when a suitable consist fleet size is available, savings strongly concentrate on (long travel) intermodal trains and that to obtain significant savings only a small number of consist types changes are sufficient. The future studies should include other homogeneity parameters in the implementation of the consists selection. Along with the fueling homogeneity we suggest to build consists considering the maintenance homogeneity too. According to the U.S. Federal Railroad Administration requirements, each locomotive must undergo preemptive maintenance at some designated shop before 92 days have elapsed since its last maintenance. Thereby, locomotives are sent to maintenance centers (shops) every 92 days (Nahapetyan et al. [2007], Illés et al. [2006]), and a locomotive becomes critical when its maintenance is scheduled within 7 days. In general, the residual time to the next maintenance event (hereinafter \( RTM \)) is different for each locomotive inside each consist, thereby consists are in general heterogeneous with respect to the \( RTM \) parameter. This fact has two important consequences in the routing phase:

1. heterogeneous consists are busted to maintain critical locomotives;
2. critical locomotives are highly dispersed over many different stations;

The \( RTM \) heterogeneity may increase consist bustings (needed to send critical locomotives to the shops) and the dispersion of critical locomotives, and may increase:

1. the number of travels toward the shops \( \Rightarrow \) high travel costs;
2. the organizational/logistic complexity;
3. the operational risks for crews and equipment.

Building consists considering the \( RTM \) parameter permits to obtain homogeneous consist with the following positive impacts in the routing phase:

1. critical locomotives are grouped in critical consists, minimizing the number of stations where critical locomotives are located (low dispersion) and the travels towards shops;
2. a critical consist may be sent to the shop in its entirety, thereby avoiding a busting operation.

The trivial example depicted in Figure 15 exemplifies the reduction of both travels toward shops and consist busting.
Working with locomotives grouped in RTM homogeneous consists could require an increased locomotive fleet size to guarantee the turnover and preserve the feasibility of the weekly plan. Contrary to the fueling homogeneity, the maintenance homogeneity policy affects the number of needed locomotives. For this reason, to obtain an evaluation of the cost-benefit ratio of the maintenance homogeneity policy, future studies should solve the consists selection jointly with the LPP optimization.

References


ENVIRON International Corporation. Port of [o]akland 2005 seaport air emissions inventory. Tech report,


W. Sylte. Port of Oakland 2005 seaport air emission inventory. Tech report, ENVIRON International Cor-


