Model for Management of Empty Freight Cars

VEENA B. MENDIRATTA AND MARK A. TURNQUIST

A model for the distribution of empty freight cars is presented. The perspective on empty-car distribution developed in this paper views the problem as a lack of coordination between decisions made centrally for the railroad as a whole and decisions made locally at individual terminals. Consequently, the model incorporates interacting submodels—the network model and the terminal model—that represent activities performed by these two levels in the railroad system. The objective of the network model is to maximize profits when the entire network, subject to the constraints imposed by empty-car supply and demand. The terminal model is an inventory-control model that incorporates stochastic demands and lead times for delivery of empty cars. Coordination between the two levels is achieved through internal transfer prices that reflect the opportunity costs of cars. The model can be used as a policy evaluation tool by railroad central management and as an operational tool for the daily distribution of empty cars by terminal personnel. Tests of the model by using data from a cooperating railroad indicate that it leads to distribution decisions that reduce empty-car miles, empty trips, and empty-car days without reducing the percentage of demand satisfied.

Rolling stock represents one of the largest capital investments by most railroads but it is an asset that is very poorly used in general (1). For many years, the railroad industry has had difficulty maintaining an adequate return on investment. The cost of owning and operating freight cars has been rising rapidly, and continued low levels of use certainly contribute to the financial problems of the industry. The focus of the research in this paper is to improve the empty-car distribution process in order to achieve higher levels of car use. Empty-car distribution is the process of controlling the flow of empty cars from the time they are unloaded to the time they are placed for the next load or delivered off-line.

On many railroads, the primary responsibility for empty-car distribution is at a relatively local level. Often, major decisions are made by managers at individual terminals. Poor service reliability, fluctuations in car supply and demand, and the absence of terminal management accountability for the costs of empty-car time and movement lead to ineffective car management. In times of car shortage, overstocking of empty cars at terminals that terminate more traffic than they originate is observed. This is an important symptom of a general problem. Terminal managers are faced on the one hand with a need to satisfy uncertain shipper demands for empty cars. They perceive very clearly a cost (at least a subjective cost) of not being able to meet those demands. On the other hand, they generally are not financially accountable for the time and cost of cars held in their yards. In such a situation, it is quite understandable that terminal managers hold more cars than necessary, and this tends to reduce the effectiveness of car distribution and degrade car use.

In times of car surplus, many terminals have more empty cars than they need and tend to dispatch them to another location to "get them out of their way." This leads to excessive empty-car mileage.

An important element of a solution to this problem is to make terminal managers more aware of the opportunity costs of decisions made on distributing empty cars. In particular, the opportunity cost of holding an empty car is the cost of foregoing the opportunity to use the car somewhere else in order to hold it where it is currently. In times of car surplus, this cost may be substantial. Conversely, when cars are in short supply, this opportunity cost may be substantial. If it is larger than the movement cost of repositioning the empty car, it should be moved to accept a new load. The opportunity cost reflects the fact that the value of an empty car is a function of when and where it is available for use. If the opportunity costs of empty-car time appeared in the accounts of the terminal, there would be an incentive for terminal managers to use this resource more efficiently.

The determination of the opportunity costs for empty cars must be done at the central, or corporate, level. This is because the value of empty cars depends on activity of the entire network. However, complete centralization of empty-car distribution decisions is not always desirable. These decisions must be communicated to the terminals and then implemented at that level. The implementation of car-distribution decisions is likely to be more effective if terminal managers play an active role in the decisionmaking process.

MODEL OVERVIEW

The approach taken in this paper is to (a) recognize that certain decisions can be made at each level (corporate and terminal) and (b) provide a mechanism for coordinating those decisions. This has been accomplished by creating a system model that incorporates interacting submodels that represent activities at two levels in the railroad: (a) the central decisionmaking at corporate headquarters that concerns movements over the total railroad network (the network model), and (b) the inventory-sizing decisions at individual railroad terminals (the terminal model). The objective of the system model is to maximize profits for the railroad subject to the constraints imposed by available empty cars, shipper demands for empty cars, and institutional requirements. An important aspect of the model is that it functions even when there is an overall shortage of empty cars in the system.

In order to achieve system optimality, there is an iterative exchange of information between the network model and the terminal model. The network model determines internal transfer prices for empty cars that are input to the terminal model; the terminal model uses these prices to determine orders or releases of cars. These orders and releases are input to the network model for recomputation of transfer prices. This cycle continues until the results of the two component models are consistent.

Optimality, in this context, means that the transfer prices produced by the network model result in a pattern of orders and releases among the various terminals in the network that, when input to the network model, reproduce the same prices. The solution is optimal in the sense that it maximizes net contribution to profit, given the various input parameters to the model system as a whole.

If this type of model were implemented, the transfer prices for empty cars produced at the optimal solution would provide one important input to the decisions of terminal managers regarding the ordering or release of empty cars. The distinction between the terminal model and the terminal manager must be emphasized. The model is intended to be an aid to decisionmaking at the terminal, not to replace the manager. The person who makes the decisions could consider aspects of the situation not reflected in the model and might choose to modify the recommendations of the model. This is another advantage of the methodology developed: The manager can examine the effects on costs of his or her
modifications. By feeding the modifications back into the model, the effect on the individual terminal and on the rest of the system can be studied.

In addition to transfer prices, the network model produces a pattern of empty-car flows (i.e., where to distribute cars) over the network. When the optimal solution has been reached, the flows provide a basis for car distributors' decisions to reposition empty cars. They also could be modified as necessary to meet special conditions not reflected in the model. Again, the model provides the ability to trace out the costs and effects of modifications.

Figure 1 illustrates the structure of the model and how the model results could be used by railroad terminal management and car distributors. The following sections describe the network and terminal models in more detail.

NETWORK MODEL

In the context of the empty-car distribution problem, the divisions are interdependent because the empty cars at the supply terminals are used to meet orders at the demand terminals. The network model is a price-directed decentralized resource-allocation procedure. It provides a mechanism for the central railroad management to determine transfer or shadow prices for empty cars. Mathematically, the network model can be formulated as a decomposable optimization problem with a structure similar to that of the so-called transportation problem.

There is a substantial history of formulating empty-car distribution problems by using network-optimization models based on the transportation problem (2-4). These models have met with limited success. The most obvious reasons for this are that they are static and deterministic models, they impose complete centralization of decisionmaking, and they do not function well when cars are in short supply. By incorporating appropriate price information and by constructing the network model to emphasize decomposition, we are able to overcome the last two of these difficulties. The incorporation of stochastic and dynamic effects is an important aspect of the terminal model, which is discussed in the next section.

The network model functions in the following way. Headquarters (the master problem) determines tentative prices for the common resources (empty cars) and transmits the prices to the divisions (subproblems). Each subproblem determines the desired number of empty cars for a particular terminal at the given prices and transmits that information to the master problem. The prices are then adjusted to reduce prices of the resources in excess supply and increase the prices of resources for which there is excess demand. This iterative exchange of information is continued until an optimal solution is found.

The objective of the network problem is to minimize the sum of transportation, inventory holding, and shortage costs over all terminals. The transportation costs refer to the cost of moving empty cars between supply and demand nodes. The shortage costs are the opportunity costs of unsatisfied demand at a node. In the long run, the inventory holding costs would equal the ownership costs of the equipment and would depend on the interest rate, depreciation rate, and market price of a car. However, in the short term, the overall fleet size is fixed. Therefore, the cost of holding an empty car on a given day is the marginal value (shadow price) of the car at its current location. The shadow price is a function of the overall supply and demand for cars on that day and of the locations of those supplies and demands.

Given the variable costs of empty-car distribution, the network problem may be stated as follows:

\[
\text{minimize} \quad \left[ \sum_{j=1}^{n} - \sum_{i=1}^{n} C_{ij} X_{ij} + \sum_{j=1}^{n} D_{j} \right] \\
\text{subject to} \\
\sum_{j=1}^{n} X_{ij} \leq D_{i} \quad j = 1, 2, \ldots, n \\
\sum_{i=1}^{n} X_{ij} \leq S_{i} \quad i = 1, 2, \ldots, n \\
X_{ij} \geq 0 \quad \text{for all i-j pairs}
\]

where

- \(X_{ij}\) = flow of empty cars from node \(i\) to node \(j\),
- \(C_{ij}\) = unit transportation cost from node \(i\) to node \(j\),
- \(R_{j}\) = revenues per carload that originate at node \(j\) less the transportation cost of the loaded trip from node \(j\),
- \(D_{j}\) = gross demand for empty cars at node \(j\), and
- \(S_{i}\) = gross supply of empty cars at node \(i\).
Note that each terminal is potentially both a supply node and a demand node and that
\[ (D_i - \sum_{j=1}^n X_{ij}) = \text{unmet demand at node } i, \]
and
\[ (S_j - \sum_{i=1}^n X_{ij}) = \text{surplus cars at node } i. \]

The model, as formulated above, is capable of solving the problem for one car type. Alternatively, the model can easily be formulated as a multicommodity-type optimization problem by including the supply and demand constraints for each car type and capacity constraints for each i-j link in the rail network. Capacity constraints at a yard can be incorporated in the model by adding links to the network that represent the yard.

It is assumed that a demand node cannot use any cars in excess of its requirements. If demand node \( j \) receives fewer cars than it demands, it incurs a shortage cost equal to \( R_j \).

The opportunity cost of holding cars is the dual variable associated with each of the supply constraints (Equation 3). If there is a surplus of cars at node \( i \), the corresponding constraint is slack and the dual variable, or shadow price, is zero. This implies that in times of surplus, the marginal cost of holding a car is zero.

By deleting the constant terms and by changing the sign of the objective function, the network problem may be rewritten as follows:

\[
\text{maximize } \sum_{i=1}^n \sum_{j=1}^n (S_j - C_{ij})X_{ij}
\]
subject to
\[
\sum_{i=1}^n X_{ij} < D_j, \quad j = 1, 2, \ldots, n
\]
\[
\sum_{j=1}^n X_{ij} < S_i, \quad i = 1, 2, \ldots, n
\]
\[
X_{ij} > 0 \quad \text{for all } i-j \text{ pairs}
\]

This problem is similar in structure to a transportation problem without the requirement that
\[
\sum_{j=1}^n D_j < \sum_{i=1}^n S_i.
\]

Consequently, the inequality in the set of demand constraints (Equation 6) is "less than or equal to". This implies that it is not necessary to satisfy all the demand in the system even when empty cars may be available. Demand will be satisfied only when it is profitable to do so; that is, when revenue exceeds variable costs. This can also be seen from the coefficients of the objective function. Since the objective function is being maximized, \( X_{ij} > 0 \) only if \( R_j > C_{ij} \); that is, only if the net revenue from loading a car at node \( j \) is greater than the transportation cost from node \( i \) to node \( j \).

The problem expressed in Equations 5-8 may be decomposed by divisions. The subproblem to be solved by each division \( j \) is given as follows:

\[
\text{maximize } \sum_{i=1}^n (R_i - C_{ij} - U_j)X_{ij}
\]
subject to
\[

\sum_{i=1}^n X_{ij} < D_j
\]
\[
X_{ij} > 0
\]
where \( U_1 \) is the dual variable that corresponds to the \( i \)th supply constraint in Equation 7 and is the transfer price for cars at terminal \( i \) computed by the master problem.

The linearity of the divisional subproblems implies that their solutions will be at extreme points of their feasible sets. To ensure an overall system optimal solution, the master problem (headquarters) must create weights to apply to the various subproblem solutions (5, 6).

The point to be emphasized here is that headquarters actually makes the final decision by issuing orders to the terminals on the number of cars to be shipped or ordered. From an organizational point of view, this breakdown of the decentralized decisionmaking structure presents problems of enforcement and motivation of division managers. These problems can be overcome by modifying the nature of the decomposition model.

Jennergren (7) has developed a method that allows the divisions (terminals) to make their own calculations of resource requirements that are both locally optimal and system optimal. This allows lower-level managers to retain decisionmaking responsibility rather than simply being allocated resources by central management (headquarters). The basic idea of this method is to supply terminals with a set of transfer price schedules instead of resource allocations. The price schedule given to subproblem (terminal) \( j \) for cars from terminal \( i \) is of the following form:

\[
P_{ij} = U_j^* - 2X_{ij} + kZ_{ij}
\]

where
\[
P_{ij} = \text{price at which terminal } j \text{ may order empty cars from terminal } i,
\]
\[
x_{ij} = \text{number of empty cars ordered (} x_{ij} > 0 \text{)}
\]
\[
by \text{ terminal } j \text{ from terminal } i \text{ or delivered (} x_{ij} < 0 \text{) to terminal } i,
\]
\[
x_{ij}^* = \text{optimal solution from problem in Equations 5-8 for a given set of } D_j \text{ and } S_i \text{ values},
\]
\[
U_j^* = \text{optimal dual variables (shadow prices)}
\]

The derivation of the form of these price schedules may be found in Mendiratta (8).

Note that because the price schedule is a function of the number of cars ordered or released by the terminal, the individual subproblems become quadratic rather than linear. The structure of these modified subproblems is discussed in the next section.

Operationally, the model works as follows. An initial set of \( S_i \) and \( D_j \) values is assumed, and the linear program in Equations 5-8 is solved. This yields a solution vector of \( X_{ij}^* \) and dual variables \( U_j^* \). These values are used to specify \( P_{ij} \) as a function of \( Z_{ij}^* \). This function is used in the subproblems (terminal model) to compute the optimal \( x_{ij}^* \) (\( = I \times Z_{ij}^* \)) for each terminal \( j \).

If \( Z_{ij}^* < 0 \), node \( j \) is a net supply node, and we set \( S_j = -Z_{ij}^* > 0 \), it is a net demand node, and we set \( D_j = Z_{ij}^* \). The set of \( S \) and \( D \) values are then input to the network model for the next iteration, unless they are essentially unchanged from the
previous iteration, in which case the process is terminated and the optimal solution reported.

The exact mechanism by which the terminal model uses the \( P_{ij} \) functions to compute \( z_j^* \) is discussed in the following section.

**TERMINAL MODEL**

Inventory-sizing decisions by terminal managers with respect to empty cars impact all aspects of railroad operations. Each day the terminal manager makes a decision on the number of cars to order or release to other terminals. A related question is when to place the order so that future demands are satisfied. The inventory-sizing problem at a terminal is different from the standard inventory-control problem in at least one important aspect: A terminal may have a surplus or deficit of empty cars on any given day and the terminal has the option of disposing or ordering cars to correct the situation. This has been a major consideration in developing the terminal model, which is basically an inventory-control model that incorporates stochastic demands and lead times for delivery of cars. In its basic concept, this model is similar to that developed by Philip and Sussman (9), although the details of the formulation are substantially different.

The inventory-control problem at a railroad terminal, with respect to the distribution of empty cars, requires determining the number of cars to order at a demand node or the number of cars to dispatch from a supply node. Inputs to the system are the supply of empty cars available at the terminal, the demand for empty cars at the terminal, outstanding orders of empty cars, expected arrival time of outstanding orders, delivery time associated with orders, and car holding, shortage, and buying or selling costs.

In general, the inventory balance equation is of the form:

\[
x(t+1) = x(t) + x_d(t) - x_r(t)
\]

where

- \( x(t+1) \) = net inventory at beginning of time period \( t \) or at end of time period \( t \)
- \( x_d(t) \) = cars received or released by shippers during time period \( t \) from previous orders, and
- \( x_r(t) \) = cars dispatched during time period \( t \) to shippers or to other terminals.

The objective of an inventory-control model is to determine an operating policy or rule for the decision variable \( z(t) \) (number of cars to order or dispatch in time period \( t \)), such that the total expected costs are minimized over the planning period. Only costs that vary with the operating rule are included in the objective function. The relevant costs for the empty-car distribution model are described below.

**Holding Cost**

The holding cost (in dollars per car per day) is the cost of carrying inventory. In the long run, it is the ownership cost of equipment, a function of the replacement cost of a car, the interest rate, and the equipment depreciation rate. In the short run (the case of interest to us), the amount of equipment is fixed and it is the opportunity cost of holding a car at a terminal. This cost is the shadow price \( U_i^* \) associated with the supply constraints of the network model described above.

**Shortage Cost**

The shortage cost (in dollars per car per day) is the cost incurred by a terminal if a demand occurs when the terminal is out of stock of empty cars. Unsatisfied demands are assumed to be back ordered. The shortage cost per day at terminal \( j \), \( C_{ij}^s \), is approximated by

\[
C_{ij}^s = \left[ (R_j + W) \Delta \right] / 365
\]

where

- \( R_j \) = net revenues per carload (revenue less the cost of the loaded trip) that originates at terminal \( j \)
- \( W_j \) = value per carload of the goods to be shipped, and
- \( \Delta \) = annual interest rate on capital.

\( R_j \) and \( W_j \) can be adjusted periodically, depending on the type of commodities shipped, to account for the changing mix of shipments from a terminal.

**Order/Release Cost**

The major components of the order/release cost (in dollars per car) are the cost of ordering (or releasing) the car, computed from the price schedule described above, and the transportation cost \( C_{ij} \).

There are other cost components, such as the costs of actually processing the order. However, these are relatively minor and are not included in the model.

The purpose of developing a mathematical model of the inventory system is to determine an optimal operating rule. The criterion for selecting the operating rule is the minimization of costs summed over the planning period. The costs in time period \( t \) are a function of the net inventory at the end of the period \( x(t+1) \) and the number of cars ordered or dispatched in the period \( z(t) \). If \( f(x(t+1), z(t)) \) denotes the costs of holding inventory or back orders and buying or selling cars in time period \( t \), the problem is to determine an optimal value for \( z(t) \), \( t = 1, \ldots, T \), such that the performance criterion is minimized:

\[
J = \min \left\{ E \left[ f(x(t+1), z(t)) \right] \right\}
\]

where

- \( T \) = planning period over which \( J \) is computed,
- \( f \) = function of \( x(t+1) \) and \( z(t) \) derived from the car holding, shortage, and order/release costs; and
- \( E(\cdot) \) = expected value.

The expected value of costs is minimized because, in general, \( x(t) \) is not deterministic.

If it is assumed that the performance criterion includes linear and quadratic terms for the state variables \( x(t) \) and control variables \( z(t) \), then it is possible to formulate the inventory problem as an optimal-control problem and use standard results from optimal-control theory. For problems of this type the optimal operating rule is a continuous linear function of the state variable \( x(t) \).

For the case where first- and second-order terms are included in the performance criterion, the optimal-control problem for a linear stochastic system is given by Kleindorfer and Kleindorfer (11):

\[
z(t) = -G(x(t)) - (1/2)H^{-1}(t)C'(t)
\]
where $X(t)$ equals $E[X(t) | U_t]$ is the best estimate of state $X(t)$ at the $t$th instant and is the vector of net inventory of empty cars and empty cars in transit with respect to a terminal; and $G(t)$, $H^{-1}(t)$, and $C'(t)$ are the output of the optimal feedback control algorithm.

The optimal feedback control algorithm for stochastic linear systems is a set of recursive relations for solving an optimal-control problem for a given time period by using dynamic programming. For a more complete description of the algorithm and its application to the inventory-sizing problem at terminals, see Mendiratta (3).

The assumption of a quadratic performance index is not restrictive because any cost function can be approximated by a second-order Taylor-series expansion to yield the desired form. The advantage of a model of this type is that the same model is applicable to both supply and demand nodes. The model output, $x^0(t)$, is returned to the network model as the net supply or demand for this terminal in the given time period.

MODEL TESTING

The empty-car distribution model has been tested in two ways. First, a series of experiments have been conducted by using hypothetical data to observe the sensitivity of the model to various input parameters. A second form of testing has been to compare model output with observed performance on an existing railroad system. The following measures are the basis for analyzing the results of the experiments:

1. Number of empty trips,
2. Number of empty-car miles,
3. Number of empty-car days, and
4. Percentage of demand satisfied on time.

A comparison of test results with the observed performance at a railroad indicates that model performance is much superior with respect to minimizing empty-car miles, empty trips, and empty-car days—all important measures of car use. With respect to demand satisfied, the performance of the model is comparable to performance in the real system. These results are shown graphically in Figure 2. The implication is that equivalent service quality could be provided at much lower cost through use of the methods for car distribution described in this paper.

The model is not overly sensitive to changes in either the price schedule parameter or the interest rate used in computing the shortage costs due to unsatisfied demand. We recognize, of course, that the improvements attainable in an operational setting may not be as dramatic as these test results suggest. Running an optimization model with observed data after the fact should generate better results than those attained in practice. Day-to-day decisions must be made with partial and imperfect information, and hindsight inevitably is sharper than foresight. Nevertheless, the experimental results certainly indicate the potential usefulness of this model as an aid in making car distribution decisions.

USES OF THE MODEL

The ideas embodied in the empty-car distribution model could be used by a railroad in two different ways: simulation (test and evaluation of potential improvements) and operational use.

Simulation of Empty-Car Movements

In the simulation of empty-car movements, the model would provide both an evaluative tool for management to assess the effectiveness of current procedures and a planning tool to help design improved distribution procedures. This use of the model requires that central management obtain the necessary inputs and operate all the components of the model. Over time, management could compare the output of the model with observed empty-car movements in the system. Model output could also be used to develop operating guidelines for terminal personnel.

Operational Implementation

In using operational implementation, central management would operate the model daily to obtain opportunity costs, or transfer prices, for cars at each terminal. This information (the price charged to the terminal for the use, or ordering, of an empty car on that day) would be sent to each terminal. Similarly, a terminal with surplus cars would receive a credit for each car that it dispatched to another terminal. The terminal manager's decision on ordering or dispatching cars would be made on the basis of the current price, in addition to other information currently used, such as shipper demand.
expected supply of empty cars, and backlog of demand. Central management could also make available to the terminals the operating guidelines developed by using the model as a simulation tool. A very important by-product of operational implementation would be the daily documentation of the costs of system shortages and surpluses. This could serve as input to decisions regarding disposal or acquisition of cars in the long term. By using this information, a railroad could develop a relation between fleet size and the volume of demand that can be satisfied over a given network and flow pattern.

CONCLUSIONS
An optimization model for the management of empty freight cars has been developed. The perspective on which this research is based is that an important element of the car-distribution problem is the lack of coordination between decisions made centrally for the railroad as a whole and decisions made locally at individual terminals. The model provides a mechanism to coordinate the decisions made at these two levels. The basis of the model is the creation of internal transfer prices that reflect the opportunity costs of cars at various points in the network in each time period. The model can be used by a railroad as a simulation tool for testing and evaluating potential improvements in car-distribution practices. It can also be used operationally to obtain transfer prices for empty cars on a daily basis. This type of use would be consistent with a greater emphasis on individual terminals as profit centers for management purposes.

Tests of the model by using data from a cooperating railroad indicate that it leads to distribution decisions that reduce empty-car miles, empty trips, and empty-car days, without reducing the percentage of demand satisfied. Thus, the methods developed in this research can improve car use significantly and, consequently, enhance railroad profitability.

ACKNOWLEDGMENT
The work reported in this paper was supported in part by the Transportation Center at Northwestern University. This support is gratefully acknowledged.

REFERENCES

Where Does the Energy Go? A Simplified Perspective on Fuel Efficiency in Rail Freight Transportation

JOHN B. HOPKINS

The potential impact of measures intended to reduce fuel use in the operation of mainline freight trains can often be bounded or approximated by relatively simple calculations. In some cases, this may be sufficient to reject options that are intuitively appealing but actually offer very small gains, thereby avoiding expensive simulation or measurement programs. When a potential conservation measure is found to have sufficient promise to warrant detailed examination, preliminary estimates can ensure that further efforts are well-founded and properly structured. In this paper, a standard train-resistance equation and industrywide aggregated data are used to assess the effect of various equipment characteristics and operating scenarios on fuel use for a baseline case (a 4700-ton, 66-car train traveling at 40 mph) in order to develop useful approximations. Emphasis is on examination of the specific physical mechanisms associated with dissipation of energy in the movement of trains. Topics considered include car weight and aerodynamics, radial trucks, locomotive idling, power-to-weight ratio, track structure, stops, operating speed and variations in speed, and coasting strategies. Quantitative estimates of fuel-consumption impacts are presented for each factor in terms of the baseline train. Extension to other cases is facilitated by presentation of data and general formulas. One noteworthy finding is that locomotive efficiency factors and train resistance account for only half of the fuel actually consumed; the major portion of the remainder appears to be dissipated in braking and, to a lesser extent, in overcoming curve resistance.

The rapid increase in fuel costs in recent years has dramatically increased the importance of minimizing