Transportation Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Survey of Optimization Models for Train Routing and Scheduling

Jean-François Cordeau, Paolo Toth, Daniele Vigo,

A Survey of Optimization Models for Train Routing and Scheduling[1]

JEAN-FRANÇOIS CORDEAU

GERAD and École Polytechnique de Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal H3T 2A7, Canada

PAOLO TOTH and DANIELE VIGO

DEIS, Università di Bologna, Viale Risorgimento, 2, 40136 Bologna, Italia

The aim of this paper is to present a survey of recent optimization models for the most commonly studied rail transportation problems. For each group of problems, we propose a classification of models and describe their important characteristics by focusing on model structure and algorithmic aspects. The review mainly concentrates on routing and scheduling problems since they represent the most important portion of the planning activities performed by railways. Routing models surveyed concern the operating policies for freight transportation and railcar fleet management, whereas scheduling models address the dispatching of trains and the assignment of locomotives and cars. A brief discussion of analytical yard and line models is also presented. The emphasis is on recent contributions, but several older yet important works are also cited.

The rail transportation industry is very rich in terms of problems that can be modeled and solved using mathematical optimization techniques. However, the related literature has experienced a slow growth and, until recently, most contributions were dealing with simplified models or small instances failing to incorporate the characteristics of real-life applications. Previous surveys by ASSAD (1980b, 1981) and HAGHANI (1987) suggest that optimization models for rail transportation were not widely used in practice and that carriers often resorted to simulation. This situation is somewhat surprising given the considerable potential savings and performance improvements that may be realized through better resource utilization. It is also contrasting with the rapid penetration of optimization methods in other fields such as air transportation (YU, 1998).

In fact, the development of optimization models for train routing and scheduling was for a long time hindered by the large size and the high difficulty of the problems studied. Important computing capabilities were needed to solve the proposed models, and even the task of collecting and organizing the relevant data required installations that very few railroads could afford. As a result, practical implementations of optimization models often had a limited success, which deterred both researchers and practitioners from pursuing the effort.

In the last decade however, a growing body of advances concerning several aspects of rail freight and passenger transportation has appeared in the operations research literature. The strong competition facing rail carriers, the privatization of many national railroads, deregulation, and the ever increasing speed of computers all motivate the use of optimization models at various levels in the organization. In addition, recently proposed models tend to exhibit an increased level of realism and to incorporate a larger variety of constraints and possibilities. In turn, this convergence of theoretical and practical standpoints results in a growing interest for optimization techniques. Hence, although simulation-based approaches are still widely used to evaluate and compare different scenarios, one witnesses a sustained development of optimization methods capable of producing high-quality solutions to complex problems within short computing times.

Problems facing rail transportation planners can

[1]Accepted by Gilbert Laporte.

Transportation Science
Vol. 32, No. 4, November 1998

0041-1655/98/3204-0380 $01.25

be grouped into a number of classes according to the facet of the organization that is concerned. The most common approach is to represent the rail transportation system as a network whose nodes represent yards or stations and whose arcs represent lines of track on which trains carry passengers or freight. One then distinguishes between local problems involving only a node or an arc of the network, and global problems involving multiple entities. Rail transportation problems can also be classified into categories according to the planning horizon considered. At the strategic level, one is mainly concerned with the acquisition or construction of durable resources that will remain active over a long period of time. The tactical level is related to medium and short term issues, and generally involves the specification of operating policies that are updated every few months. Finally, the daily tasks that are performed by taking account of the fine detail of the system belong to the operational level. This popular hierarchical approach is explained in greater detail by ASSAD (1980a), who also gives numerous examples of problems that pertain to each category.

In this paper, we intend to review most of the recent contributions dealing with train routing and scheduling with regard to both freight and passenger transportation. We will thus cover all three levels of planning but focus our attention on global problems of train management. Because of the large size and the high degree of heterogeneity that characterizes most models, we have opted for a textual description. A more involved comparison of mathematical formulations would require focusing on a much smaller subset of models.

Most reviewed models have been proposed during the last decade, although we also cite several older but important works. Apart from a few exceptions, the survey concentrates on published and easily accessible material. We have also elected to limit ourselves to contributions dealing specifically with rail transportation, even though a lot of work done in the related areas of road and air cargo transportation is certainly relevant to the rail context. Finally, the field of railway crew management will not be treated here but we instead refer the interested reader to recent work by CAPRARA et al. (1997).

The paper is organized as follows. Section 1 introduces the necessary background and definitions concerning the reviewed material. Models for train routing and train scheduling are reviewed in Sections 2 and 3, respectively. Conclusions and an account of current research trends are presented in the last section.

1. BACKGROUND AND DEFINITIONS

WE NOW GIVE a brief description of railroads and introduce some terminology that will be used throughout the text. A more detailed account of rail operations and freight transportation is presented in the book by BECKMANN, MCGUIRE, and WINSTON (1956). The authors also provide an interesting introduction to rail modeling and optimization.

The first part of the review is devoted to routing problems in the context of rail freight transportation. Demand for freight transportation is usually expressed in terms of tonnage of certain commodities to be moved from an origin to a destination. Given these demands, the railroad must establish a set of operating policies that will govern the routing of trains and freight.

For every origin–destination pair of traffic demand, the corresponding freight may be shipped either directly or indirectly. When demand is important enough, delivery delays are obviously minimized by using direct trains as opposed to sending the traffic through a sequence of links. However, when demand does not warrant the dispatching of direct trains, delays are inevitable. Either the traffic is consolidated and routed through intermediate nodes, or freight cars have to wait at the origin node until sufficient tonnage has been accumulated.

To benefit from economies of scale, trains are thus often formed by grouping cars with various commodities and having different origins and destinations. These trains operate between particular nodes of the network, called *classification yards.* At these yards, cars are separated, sorted according to their final destination, and combined to form new outbound trains. However, because the classification process requires considerable resources, cars are not reclassified at every yard on their trip from origin to destination. Instead, cars with different final destinations but sharing some initial portion of their trips are assembled into *blocks.* Cars in the same block may then pass through a series of intermediate classification yards, being separated and reclassified only after they have reached the destination of the block. The *blocking policy* specifies what blocks should be built at each yard of the network and which cars should go into each block.

In each yard, blocks are built on *classification tracks* where they await the departure of an outbound train. The list of potential blocks that may go into each outbound train is specified by the *makeup policy.* Also, when a train passes through an intermediate classification yard, it may leave or pick up blocks of cars. A block left by an inbound train is either transferred to a different train or it is broken up and its cars are reclassified. Hence, although the origin and destination of a block may correspond to those of a train, a block may also switch trains several times before reaching its final destination.

Every loaded movement on a rail network leads to a supply of empty cars at destination. Therefore, if transportation demand is unbalanced, steps must be taken to reposition empty cars and avoid their accumulation in some parts of the network where more traffic is directed. Even if traffic is balanced in the long run, this need not be the case in the short term. Repositioning empty freight cars can thus help the railroad offer better service to its customers by reducing the average time they have to wait for cars, and decrease the capital investment associated with equipment ownership. The *freight car management problem* consists of dynamically distributing empty cars in the network to improve the railroad's ability to promptly answer requests for empty cars while minimizing the costs associated with their movement.

The second part of the survey discusses models that deal with the temporal dimension of train management. Scheduling problems appear in both freight and passenger transport, albeit in slightly different forms. In the case of freight transportation, trains sometimes operate without schedules and simply depart when they have accumulated sufficient tonnage. Although this practice is still very common in North America, it is seldom seen in Europe where freight trains usually operate according to published schedules just as they do in the case of passenger transportation. When freight trains do not operate according to a schedule, potential time slots must still be assigned to them.

Although train timetabling is usually performed at the tactical level of planning, real-time operations necessitate precise synchronization of freight and passenger train movements on the lines of the physical railway network. The lines can be made of a single track, as is often the case in North America and in most developing countries, or may contain two or more tracks, as is common in Europe. To allow trains traveling in different directions on a single-track line to meet, *sidings* are located at regular intervals along the line. These short track sections allow one train to pull over and free the way for the other one. Sidings are also used to permit a fast train to pass a slower one. Given a train timetable, the *train dispatching problem* determines a feasible plan of meets and overtakes that satisfies a system of constraints on the operation of trains.

Finally, a related scheduling problem concerns the use of the rolling equipment stock. Because of the high capital expenditures associated with locomotives, a major concern to every railway is to maximize the use of these resources. The basic *locomotive assignment problem* consists of assigning a set of locomotives to cover all scheduled trains at minimum cost while satisfying some side constraints such as compatibility restrictions and maintenance requirements. Whereas freight trains generally contain a large number of cars and several engines, passenger trains use a small number of cars coupled with a few locomotives. In the case of passenger trains, it is thus possible to perform the simultaneous assignment of both types of equipment to the trains.

2. ROUTING PROBLEMS

OPERATING PLANS FOR rail freight transportation indicate the train connections to be provided, the blocks to be built in each yard, and the assignment of blocks to trains. In addition, train timetables must be developed to specify the departure and arrival times of trains. These closely intertwined policies should ideally be determined concurrently to identify the most efficient way of delivering all traffic while satisfying a set of technological constraints on train and yard capacity. However, because this leads to a very difficult problem, a sequential approach is often adopted. For example, a blocking plan may be developed first, followed by a train routing and makeup plan. Very frequently, train timetables are specified last and are designed around the routing plans. Operating plans are usually updated every few months, but weekly or daily adjustments must be made to account for demand variability.

Most optimization models for train and freight routing are defined over a network whose nodes represent origins, destinations, or intermediate transfer points for the traffic to be routed. The arcs then represent existing or possible train connections between these points that are often aggregated to represent the activities of a wider geographic area.

Because yard activities constitute an important part of freight transportation operations, we first present a brief review of analytical models developed to analyze yard performance under different configurations or traffic conditions. Although these are not optimization models per se, they may appear within the objective or constraint structure of large-scale routing models. We then present network models that address the blocking, makeup, and routing problems. Models aimed specifically at the freight car management process are described in the last section. In each section, models are presented in ascending chronological order.

2.1 Analytical Yard Models

Yard policies concern the specification of the activities to be performed in the yards of a rail network. More precisely, they indicate how trains entering each yard should be inspected and disassembled, and how cars should be sorted and reassembled into blocks that will form new outbound trains. Although reclassification work is also performed to some extent in less-than-truckload and air cargo transportation, the delays associated with these activities are usually negligible for these modes, but they constitute a large portion of the overall transit time for rail freight. As explained by KEATON (1989), car time in intermediate terminals occurs in classification and assembly operations and while waiting for the departure of an outbound train, but also as a result of yard congestion. Car time is also spent in origin and destination terminals where cars wait either for the departure of an outbound train or for delivery to the receiver by a local train.

Two types of classification yards are in common use. *Flat yards* use engines to move cars from an inbound train to classification tracks. In *hump yards*, this work is performed by gravity: cars detached from an inbound train are pushed over the top of a hump and roll down to the appropriate track. Early work on yard modeling was realized by CRANE, BROWN, and BLANCHARD (1955) who presented an analysis of a particular hump yard and discussed the queuing processes identified in inspection and classification operations. A simple model for the location of a classification yard was then proposed by MANSFIELD and WEIN (1958).

A more detailed analysis of railyard operations was performed by PETERSEN (1977a, b) who developed queuing models to represent the classification of incoming traffic and the assembly of outbound trains. In these queuing models, the basic units of arrival are complete trains to be processed. The author also modeled the delay to a railcar from the end of classification to the start of the train assembly operation with a bulk queue, and observed that this delay is a minor source of yard congestion in comparison with classification and assembly operations. The models are used to compute the probability distribution of connection times for various levels of traffic given known service times. In the second paper, expressions are derived to relate the classification and assembly times to the physical characteristics of the yard and traffic attributes. The accuracy of the models was validated using historic data from two railroads. An insightful description of railyards is also presented in the first of these papers.

TURNQUIST and DASKIN (1982) modeled yard operations from the perspective of freight cars, rather than from the perspective of trains. They thus developed queuing models for classification and connection delays that consider individual cars as the basic units of arrival. Their approach also differs from that of Petersen in the sense that connection to an outbound train and assembly are treated as a single operation. Expressions for the mean and variance of classification and connection delays are derived under the assumption of Poisson arrivals using a batch-arrival and a batch-service queuing model, respectively. The authors also demonstrated how their model may be used to evaluate the effects of train dispatching strategies on the mean and variance of delay. In particular, they analyzed two strategies that consist, respectively, of scheduling trains at regular intervals, and dispatching trains when a given number of cars become available.

A different approach to the problem of predicting yard time distributions was studied by MARTLAND (1982), who described a methodology for estimating the total connection time of cars passing through a classification yard. The model is based on a function, calibrated using actual data from the railroad, that relates the probability of making a particular train connection to the time available to make that connection and other variables such as traffic priority and volume. The function can be adjusted through different techniques such as regression analysis or simulation experiments. The approach, which has been tested and implemented by several railroads, is proposed as an aid to planning but also as a way to control operations by setting standards for train connection performance.

Other analytical models concern the performance and resource requirements of sorting strategies that specify what blocks should be assigned to each available classification track and how individual cars should be handled. Early work on this topic was performed by SIDDIQEE (1972), who compared four sorting and train formation schemes in a railroad hump yard. A screening technique and a dynamic programming approach were suggested by YAGAR, SACCOMANNO, and SHI (1983) to optimize humping and assembly operations.

DAGANZO, DOWLING, and HALL (1983) investigated the relative performance of different multistage sorting strategies. In multistage sorting, several blocks are assigned to each classification track, and cars must be resorted during train formation. Equations are derived for the service time per car of triangular sorting in both flat yards and hump yards. In a series of three papers, different classification strategies were also analyzed and compared

TABLE I
Characteristics of Network Routing Models

Authors	Problem Type	Planning Horizon	Objective Function	Model Structure	Solution Approach
Bodin et al. (1980)	Blocking	Tactical	Min operating and delay costs	Nonlinear MIP	Heuristic
Assad (1983)	Blocking	Operational	Min total classification	Shortest path	Dynamic programming
Van Dyke (1986)	Blocking	Tactical	Min operating costs	Shortest path	Heuristic
Newton (1996)	Blocking	Tactical	Min operating costs	NDP with node budget	Dantzig–Wolfe decomposition
Crainic et al. (1984)	Routing/makeup	Tactical	Min operating and delay costs	Nonlinear MIP	Heuristic decomposition
Haghani (1989)	Routing/makeup	Operational	Min operating and delay costs	Nonlinear MIP	Heuristic decomposition
Keaton (1989)	Routing/makeup	Tactical	Min operating and time costs	Linear 0–1 IP	Lagrangian relaxation
Keaton (1992)	Routing/makeup	Tactical	Min operating and time costs	Linear 0–1 IP	Lagrangian relaxation
Martinelli and Teng (1996)	Routing/makeup	Tactical	Min transit time	Nonlinear 0–1 IP	Neural networks
Marín and Salmerón (1996)	Routing/makeup	Tactical	Min operating costs	Nonlinear IP	Local search heuristics
Morlok and Peterson (1970)	Compound	Tactical	Min operating and time costs	Linear MIP	Branch-and-bound
Huntley et al. (1995)	Compound	Tactical	Min operating costs	Nonlinear MIP	Simulated annealing
Gorman (1998)	Compound	Tactical	Min operating costs	Linear 0–1 IP	Genetic search

by DAGANZO (1986, 1987a, b), who gave expressions for the switching work and space requirements. In the last two papers, the author considered dynamic blocking in which the assignment of blocks to classification tracks is allowed to vary through time.

Finally, AVRAMOVIĆ (1995) modeled the physical process of cars moving down the hump of a yard. This process is represented by a system of differential equations that incorporate several factors, such as hump profile and rolling resistance, affecting the movement of a car. The model can be used in the design of a hump yard to evaluate the strength of track retarders that regulate the speed of cars.

2.2 Network Routing Models

We now discuss network optimization models that address the different problems related to freight train routing. We first review models dealing with the blocking policy, followed by models addressing the train routing and makeup problem. Compound models that integrate blocking, makeup, and scheduling decisions are discussed last. The characteristics of the most important contributions are summarized in Table I.

2.2.1 Blocking Models

A blocking policy is usually specified as follows: cars at yard i which are destined for yard j must be added to a block that will next be shipped to yard k (possibly transiting by other intermediate yards). As explained in the introduction, cars in a block will not be reclassified until the block reaches its final destination. A blocking model thus places the emphasis on the movement of cars as opposed to the movement of trains. Its solution indicates the routing of freight through the network and the distribution of classification work among yards, but does not specify the trains to be run or the assignment of blocks to trains. Instead, an additional problem must then be solved to determine the routing of trains and their makeup.

One of the first models for car blocking belongs to BODIN et al. (1980), who suggested a nonlinear, mixed integer programming formulation of the problem. The model, which is a multicommodity flow problem with additional side constraints, simultaneously determines the optimal blocking strategies for all the classification yards in a railroad system. Besides flow equations that constitute the backbone of the model, yard capacity and block formation constraints are also considered. In particular, the model imposes upper bounds on the number of cars that may be classified and the number of blocks that may be formed in any given yard. This last constraint originates from the fact that each yard has a limited number of tracks on which blocks may be built. Block length constraints are also taken into consideration and guarantee that the number of cars in each block lies between a lower and an upper bound. Finally, *pure strategy* constraints are present. These constraints ensure that all cars in yard i destined for yard j are shipped to the same next classification yard. The objective function considered seeks to minimize the sum of shipping, processing, and delay costs. Delay costs are represented by piecewise linear functions of the flow on arcs of the network. With some manual intervention, the authors solved an instance with 33 classification yards and found a solution within 3% of a tight lower bound.

ASSAD (1983) proposed a solution approach for a problem defined on a line network composed of n yards, with traffic flowing from yard 1 to yard n. Cars are received at yard 1 in arbitrary order and must be separated as they proceed along the line to allow each successive yard to extract the traffic destined for it. Various classification strategies can be used to distribute the classification work among the yards. For the special case in which all yards have equal traffic, the author showed that the search for a solution minimizing the total work can be restricted to strategies in which traffic for yard i is separated only after previous traffic types $1, \ldots, i-1$ are already classified. When this assumption does not hold, a dynamic programming formulation of the problem leads to an efficient solution method. The author also discussed extensions to the case in which each yard is a potential source of traffic. It is shown that a dynamic programming formulation can still be used for this problem.

VAN DYKE (1986, 1988) described a heuristic blocking approach that has been tested or implemented by several large railroads. The system is based on an iterative procedure that attempts to improve an existing blocking plan by solving a series of shortest-path problems on a network whose arcs represent available blocks. Traffic is assigned to a particular block if the block is on the least cost path from the origin of the traffic to its destination. The cost of assigning traffic to a block depends on a number of factors such as block priority, traffic priority, physical rail lines traversed, and the characteristics of the origin and destination yards of the block. The solution to these problems determines the least cost distribution of traffic across a set of existing blocks. An interactive procedure allows the user to delete existing blocks or introduce additional blocks in the solution. Block capacity constraints are also taken into account by the heuristic.

Recently, a column generation algorithm was introduced by NEWTON (1996), who studied the more general network design problem (NDP) with budget constraints. This problem consists of minimizing the cost of flowing a set of commodities through a network while satisfying budget constraints on the fixed cost of the arcs used. The railroad blocking problem is transformed into this general framework by letting the nodes represent the classification yards and the arcs represent potential blocks that can be built. The fixed cost of offering direct service between two yards involves dedicating a sorting track at the origin yard. Hence, there is a separate node-budget constraint for each yard based on the number of sorting tracks available. Flow constraints are also used to restrict the total number of cars that may be sorted in each yard. The objective function minimizes the cost of delivering all commodities. Express and nonexpress traffics are treated simultaneously using priority constraints that limit the number of blocks used in delivering each commodity. The problem is solved using a branch-and-bound procedure with bounds at each node computed using Dantzig–Wolfe decomposition (DANTZIG and WOLFE, 1960). Using a labeling algorithm on an acyclic network, blocking paths with a negative reduced cost are generated for each commodity by solving a shortest path problem with a priority constraint. A rounding heuristic is also used to obtain good upper bounds. Disaggregating the bundle constraints that impose common upper bounds on the arcs of the network gives valid inequalities that strengthen the LP relaxation of the master problem. Branching is performed on the binary variables indicating whether an arc is chosen or not. Computational results were presented for instances with 150 nodes, 6000 potential arcs and 1300 commodities. Feasible solutions within a few percent of a known lower bound were found within a few hours on a workstation computer.

2.2.2 Routing and Makeup Models

Whereas blocking models indicate the routing of freight and the distribution of classification work among the yards of the network, routing and makeup models determine the routing and frequency of trains and the assignment of blocks to trains. In routing and makeup models, the blocking policy may be either determined endogenously or given as an input. These models thus produce a complete train and freight routing plan. However, because they do not provide actual departure times for the trains to be run, an additional scheduling problem must be solved at a later stage. Similar models for the service network design problem in the motor carrier industry were developed, for example, by POWELL and SHEFFI (1989).

Train formation plans are sometimes developed without regard to the concept of car blocking. For example, THOMET (1971) developed a cancellation procedure that gradually replaces direct shipments by a series of intermediate train connections to minimize operation and delay costs. A model for deciding which pairs of yards should be offered direct service to minimize total transit time of cars was also proposed by SUZUKI (1973), whereas LEBLANC (1976) suggested a network design model for strategic planning. One of the first efforts to integrate multiple components of the freight routing problem is credited to ASSAD (1980a) who proposed a multicommodity network flow model for train routing and makeup that incorporates some level of interaction between routing and yard activities.

A more complex problem was studied by CRAINIC, FERLAND, and ROUSSEAU (1984) who proposed a model and a heuristic for tactical planning. The model is a nonlinear, mixed integer, multicommodity flow problem that deals with the interactions between blocking, makeup, and train and traffic routing decisions. Traffic demand is divided into classes in which each class corresponds to an origin–destination pair, together with a commodity type. The model is based on a service network that specifies the feasible routes on which train services may be run. A set of feasible itineraries is defined for each traffic class. An itinerary specifies the train service path followed and the operations that must be performed at each intermediate stop. By selecting the best traffic distribution for each traffic class, one solves the freight routing problem as well as the blocking and makeup problems. The frequency variables associated with the possible train services provide a solution to the train routing problem. The objective function seeks to minimize the sum of operation and delay costs associated with itineraries and train services. By introducing the train service capacity constraints in the objective function, the authors obtain a modified problem for which they use a decomposition scheme that iterates between two problems until the improvement in the objective function after a complete iteration is less than a preset value. The subproblem determines the best traffic distribution for each traffic class for a given service level, whereas the master problem modifies the service frequencies to improve the solution value considering the given traffic distribution. The subproblem for each traffic class is solved using column generation and a descent algorithm. This solution methodology was explained in greater detail by CRAINIC and ROUSSEAU (1986), who presented a general framework for the design of the service network and the routing of traffic in the context of multicommodity, multimode freight transportation. The model and algorithm were tested on data from the Canadian National Railroads. The instance contained 2613 aggregated traffic classes and a service network with 415 links. Computational results indicated a significant cost reduction over the solution used by the railroad. A comparison with the simulation method used by the company was done by CRAINIC (1984). Readers interested in strategic planning are also referred to the work of CRAINIC, FLORIAN, and LÉAL (1990a).

As was properly highlighted by Haghani (1987), there exist intense interactions among the routing of trains, their makeup, their frequency, and the empty car distribution process. However, models that take all these aspects of rail transportation into consideration often get extremely complex if not simply intractable. The traditional approach has thus been to deal separately with the train routing and makeup problem and the empty car distribution problem. This obviously leads to suboptimal decisions, at both the tactical and operational levels. In an effort to counter the tendency of treating the empty car distribution problem at the operational level by assuming that routing and makeup decisions are given, HAGHANI (1989) proposed a formulation and a solution method for a combined train routing and makeup, and empty car distribution problem. The model is also dynamic and deals with temporal demand variability, providing empty car distribution decisions as well as the optimal time interval between consecutive train services between pairs of yards. To account for demand variations from period to period, each yard is replicated a certain number of times in a time–space network, depending on the period length and the horizon considered. This network has nodes representing inbound and outbound traffic for every yard in the physical network, and links representing routing, classification, delays, and deliveries. The decision variables used concern the flows of loaded cars, empty cars, and engines provided on the different links mentioned. The objective considered is to minimize the total cost defined by routing costs, classification costs, delay costs for classification and connection, and penalty costs. Penalties are imposed for carrying over the demand for empty cars and as a way to deal with boundary conditions on the shipments. Besides traditional flow conservation constraints on the loaded cars, empty cars, and engines, linking constraints ensure that the number of engines provided on each link is compatible with car routing decisions. This mixed integer model has a nonlinear objective function and linear constraints. It is solved with a heuristic decomposition approach that exploits the structure of the problem by solving an integer programming subproblem for the engine flow variables and a linear programming subproblem for the car flow variables. The algorithm was tested on a network with four nodes and five two-way links. On average, the solutions found by the heuristic were within 10% of the lower bound provided by the LP relaxation of the problem.

Keaton (1989) proposed a model and a heuristic method based on Lagrangian relaxation for the combined problem of car blocking and train routing and makeup. The model is based on a set of service networks that specify the possible train connections and blocking alternatives for each origin–destination pair. Upper limits are imposed on the number of blocks that can be formed at any terminal and on the number of cars assigned to any train. The objective function considers train costs, car time costs, and classification costs. The mixed integer programming model uses integer variables for train connections and continuous variables for car flows. By dualizing the constraints that link train variables and car flow variables into the objective function, one obtains a series of shortest path problems in the continuous variables and knapsack problems in the train variables. When ignoring train size constraints, the model can be solved efficiently with subgradient optimization and special update rules for the multipliers. Feasible solutions are improved by using a dual adjustment procedure and a greedy heuristic. A hypothetical rail network was used to generate an instance with 26 terminals and 333 origin–destination pairs. On average, solutions with duality gaps below 10% were obtained. However, when limits on train size are imposed, it becomes very hard to obtain tight lower bounds on the solution values. This model was used by KEATON (1991) to evaluate service-cost tradeoffs for carload freight traffic in the U.S. rail industry. He applied his formulation and solution method to hypothetical rail networks with variable train costs and concluded that the potential for reducing transit times by increasing train connections and frequency was rather limited.

In a subsequent paper by KEATON (1992), pure strategy constraints for blocking and maximum transit times for each origin–destination pair are also considered. The resulting formulation has only binary variables and results in a multicommodity network flow problem once the train variables are set. By dualizing the linking constraints between train and car flow variables, and constraints that place limits on train size and maximum transit time, the formulation decomposes into two easily solvable subproblems. In fact, a further relaxation is obtained by discarding all constraints on train size, yard volumes, and service levels, and dualizing the linking constraints between train and car flow variables. This relaxation can be solved efficiently using a dual adjustment procedure, and tight lower bounds can be generated. By iteratively solving this relaxation and adjusting the car or train costs in each iteration, a feasible solution to the original problem is finally obtained. However, this approach, called *iterative strategy*, does not yield explicit lower bounds on the cost of the original problem, and thus the quality of the solution obtained cannot be evaluated precisely. Computational experiments were performed on a set of three rail systems containing about 80 terminals and 1300 to 1500 origin–destination pairs.

Neural networks were used by MARTINELLI and TENG (1996) to solve a train formation problem. For a given distribution of demand, expressed as the number of cars to be moved between each origin–destination pair, the problem is to assign each class of demand to a unique itinerary chosen from a predefined subset. An itinerary specifies a succession of intermediate yards together with the train sequence used. The problem is formulated as a 0–1 integer program with a nonlinear objective function that minimizes the total time spent by cars in the system. A back-propagation neural network model trained with two groups of patterns was used to solve small instances of the problem. Good performance was obtained, as measured by the quality of the solutions, but the computation times were rather long. The data used contained 30 demand classes, 44 trains, and 108 combinations of demand-train assignments.

In a series of two papers, MARÍN and SALMERÓN (1996a, b) proposed and analyzed the expected performance of local search heuristics for the tactical planning of rail freight networks. Again, the model is based on a service network and considers demands given in terms of origin and destination yards and freight type. Each train service is defined by an origin yard, a set of intermediate yards, a destination yard, and technical characteristics such as speed and capacity. The objective is to minimize car costs, train costs, and investment costs incurred when not enough trains are available. This last term, which uses a crude approximation of the required fleet size, makes the objective function piecewise linear. Because each train service specifies the set of intermediate stations, restrictions on the number of cars transiting in any yard can be imposed. Constraints are also imposed on the number of cars assigned to each service given the chosen service frequency. The three heuristic methods proposed (descent method, simulated annealing, and tabu search) share a common decomposition that separates the routing of the freight cars and the choice of train service frequencies. The first subproblem, which is solved through a sequential loading algorithm, determines the best routes for a given choice of train frequencies. The second subproblem, which may be solved by inspection, readjusts the train frequencies for the given car routing. In each iteration of the various heuristics, train frequencies are updated according to a move that is chosen from the neighborhood of the current solution, and the car routing subproblem is solved. A reformulation of the problem as a linear program leads to an exact branch-and-bound algorithm that can be used for comparison purposes with the heuristics. Computational tests on four generated networks showed that simulated annealing obtained the best solutions but required more time than the other heuristics. This conclusion was also confirmed by the statistical analysis conducted in the second paper. The largest instance solved contained 82 train services and 150 demand classes.

2.2.3 Compound Routing and Scheduling Models

Routing and makeup models produce a transportation plan that completely describes the routing of freight, the set of trains to be operated, and their respective frequency. But because these models do not take scheduling into consideration, it may be difficult to later find a timetable accommodating all planned trains and satisfying line and yard capacity. Hence, compound models, which address both the routing and the scheduling aspects of freight transportation, can significantly help to improve service reliability and reduce costs. The recent work of FARVOLDEN and POWELL (1994) described a similar approach for the motor carrier industry. Also, railroad revenue management models based on profit maximizing and load selection formulations were introduced by CAMPBELL (1996) and KRAFT (1998).

One of the first efforts to integrate both routing and scheduling decisions into a single optimization model is probably the work of MORLOK and PETERSON (1970). Given a network representing the possible train connections, a binary variable is associated with each train service that may be operated. Each such service is defined by a route in the network, a set of stops, a departure time at the initial node, and additional attributes such as speed and capacity. A second set of binary variables is used to represent the assignment of demand to trains. Additional variables are also introduced to keep track of car time in the network. The costs considered include train and engine crew costs, intermediate yard costs, and car time costs. Besides traditional demand constraints, the model incorporates constraints on the maximum number of cars per train as well as scheduling constraints requiring that certain cars be delivered to given yards before a cut-off time. The model was applied to a very small instance and solved with a branch-and-bound procedure.

A computerized routing and scheduling system was developed by HUNTLEY et al. (1995) to help planners at CSX Transportation account for the effects of routing and scheduling decisions in strategic planning. Demand is represented as batches that have associated origin and destination yards. Each pair of switching yards in the network defines a link that may accommodate a certain number of trains. The output of the model is the sequence of train links that each batch should follow from origin to destination, as well as the departure times for all train links. The nonlinear objective function minimizes operational costs defined by fuel cost, crew cost, locomotive capital cost, and freight car rental cost. The problem is solved using simulated annealing and a perturbation operator that inserts or deletes a stop from the route of a batch, and adjusts the departure times of the trains. The system was tested on a real problem involving 166 batches and 41 yards. Related field testing showed that the system was useful in analyzing a variety of scenarios, and produced schedules having similar properties to those of the solutions in use by the company, but a smaller cost.

A combination of genetic and tabu search algorithms were used by GORMAN (1998) to address the weekly routing and scheduling problem. To solve the problem for actual train departure times, the time horizon is discretized in hours. Each train may also operate at different speeds and perform a variable sequence of stops on its way from origin to destination. The mathematical formulation has binary variables associated with each potential train service that may be operated during the week. Each possible assignment of demand to a train is also represented by a binary variable. Constraints are imposed on train size to ensure that trains operate on schedule. There are also linking constraints to enforce yard and line capacity. The objective function minimizes the sum of fixed costs of trains and marginal cost per car. The model decomposes into train-scheduling and traffic-assignment components. To solve the problem, the author suggested a classical genetic search procedure in which the population is formed by all possible train schedules. Every time an individual is generated, its cost is evaluated by solving the traffic-assignment problem. Mutations are obtained by either adding or deleting a train, or by shifting a train to an earlier or a later time in the schedule. To improve the performance of the genetic algorithm, each solution is cloned and modified with a tabu search algorithm, thus simulating the use of knowledge-based mutation operators. Computational experiments on data from a major U.S. freight railroad produced solutions that satisfied more con-

TABLE II
Characteristics of Freight Car Management Models

Authors	Problem Type	Planning Horizon	Objective Function	Model Structure	Solution Approach
Beaujon and Turnquist (1991)	Single railroad	Tactical	Max expected profits	Nonlinear network	Frank–Wolfe
Morin (1993)	Single railroad	Operational	Min operating costs	Multicommodity	Decomposition
Spieckermann and Voss (1995)	Single railroad	Operational	Min transport costs	Job-shop scheduling	Greedy heuristic
Holmberg et al. (1996)	Single railroad	Operational	Min transport and shortage costs	Multicommodity	Branch-and-bound
Adamidou et al. (1993)	Multiple railroads	Tactical	Max profits	Nash equilibrium	Gauss–Seidel
Sherali and Tuncbilek (1997)	Multiple railroads	Strategic	Min fleet size	Network	Heuristic decomposition

straints and had a smaller cost than the solution actually used by the railroad.

2.3 Freight Car Management Models

The utilization cycle of a freight car starts when a client issues an order for empty cars. At a nearby yard, compatible cars are selected and moved to a loading point. Once loaded, they are taken to a classification yard where they are sorted, assembled into blocks, and put onto outbound trains. When a car has reached its final destination, it is unloaded and, unless it is needed by the receiver, it is returned to the railroad. At this point, the car is available for a new shipment and the cycle may repeat. Very often, however, it will travel empty to a different location where a request must be fulfilled. Because demand for transportation is rarely known long in advance, the railroad must anticipate future requests and manage its fleet accordingly. A good repositioning strategy helps to reduce the size of the fleet and to decrease the delays in delivering empty cars to customers.

Models for fleet management and distribution of empty vehicles were reviewed by DEJAX and CRAINIC (1987). The management of empty railcars shares several characteristics with the distribution of empty containers used in land, maritime, or multimode transportation. Dynamic and stochastic models for the land distribution of empty containers were developed by CRAINIC, GENDREAU, and DEJAX (1990b, 1993). Also, recent work on operations planning in intermodal transportation was performed by NOZICK and MORLOK (1997). Finally, the related problem of dynamic vehicle allocation was initially studied by POWELL (1986, 1987), and later developments have been summarized by POWELL, JAILLET, and ODONI (1995).

We now review optimization models for the distribution of empty rail cars. We first discuss models used in the case of a single railroad, followed by models for the case of multiple railroads sharing a fleet of cars under a pooling agreement. The characteristics of the most recent models in each category are summarized in Table II.

2.3.1 Single Railroad Models

In the first attempts to optimize the distribution of empty freight cars, the process was often represented as a simple network flow problem for which efficient algorithms were available. WHITE and BOMBERAULT (1969) generated a network from a time–space diagram and solved the resulting transshipment problem with a modified out-of-kilter algorithm (FORD and FULKERSON, 1962). Also, static formulations solvable as transportation problems were proposed by ALLMAN (1972) and MISRA (1972). HERREN (1973, 1977) formulated a more complex problem, with a heterogeneous fleet of cars and substitution possibilities, as a minimum cost network flow model that could be solved with a specialized algorithm.

A different approach to freight car management consists of representing the system with an inventory model. One of the first efforts in this direction is the work of AVI-ITZHAK, BENN, and POWELL (1967) who suggested mathematical models for describing the behavior of car pool systems. PHILIP and SUSSMAN (1977) proposed a discrete event simulation model to determine the optimum inventory level for a single terminal. The inventory management approach was later extended to an entire network by MENDIRATTA and TURNQUIST (1982) who developed a linear programming formulation solvable by a decomposition algorithm.

One of the first contributions dealing with the stochastic nature of the problem is from JORDAN and TURNQUIST (1983), who presented a dynamic network optimization model, based on earlier work by COOPER and LEBLANC (1977), that takes into account variability in empty car demand and supply, as well as uncertainty in travel times. A methodology based on a combination of linear programming and simulation techniques was proposed by RATCLIFFE, VINOD, and SPARROW (1984) to optimize freight car dispatching given known and anticipated demands. Also, a real-life application of linear programming techniques to the daily distribution problem was presented by MARKOWICZ and TURNQUIST (1990).

A combined model for fleet sizing and vehicle distribution and assignment was described by BEAUJON and TURNQUIST (1991). Their approach takes into account the dynamic nature of these decisions as well as the uncertainty in demand and transit times. They first proposed an exact formulation, which can be viewed as a stochastic programming problem or as a stochastic control problem. Because this formulation appears computationally unattractive, a solution method was developed for an approximate reformulation of the problem. The reformulation replaces random variables associated with transportation demand and travel times by their expected value to obtain a network optimization model. The objective function maximizes the expected profit, which is defined by the difference between revenues generated by serving demands and costs incurred for vehicle ownership, vehicle movement, and unmet demand. To appropriately model the cost structure of the problem, the concept of *net vehicle pool* is introduced. At each terminal, this quantity represents both the expected vehicle pool and the expected vehicle shortage. Nonlinear costs on the arcs are then used to account for vehicle holding and unmet demand. Because the random travel times are replaced by their expectation, the network approximation introduces an error in representing vehicle arrivals. The solution procedure presented tries to circumvent this weakness by solving a pure network formulation to determine empty vehicle dispatching decisions, and adjusting the size of the net vehicle pools to account for this approximation error by solving a series of unconstrained optimization problems. The nonlinear objective includes functions of the basic decision variables that are neither convex nor concave because of variance terms. Hence, the network flow problems are solved using a procedure that iteratively fixes the variance terms, solves the resulting concave problems using the Frank–Wolfe algorithm (FRANK and WOLFE, 1956) and updates the variance terms. Numerical experiments performed on instances with up to 70 nodes and 1330 arcs showed that significant improvements are obtained by considering the stochastic nature of the problem.

Decomposition approaches were compared by MORIN (1993) who studied the empty car distribution process at SNCF and formulated the problem as a multicommodity network flow problem. Each commodity corresponds to a geographical area, and linking constraints ensure flow conservation between adjacent areas. Two formulations that can be solved with subgradient algorithms were introduced: a dual decomposition approach that relaxes the linking constraints and a primal decomposition scheme that relies on the introduction of coupling variables. The application of a mixed decomposition approach (MAHEY, 1986) that combines price-directive and resource-directive allocations was also presented with a specialized algorithm that exploits the separability of the problem. Results on a set of data from SNCF indicated that the third method was superior.

SPIECKERMANN and VOSS (1995) formulated the empty railcar distribution problem as a scheduling problem with machines representing railcars and jobs representing requests for cars. The study was realized in the context of a German car rental company that provides empty cars to its customers throughout Europe. All movements are performed by national railways to which the company must pay fees for movements of either loaded or empty cars. The objective of minimizing costs for empty moves translates into minimizing the time-dependent setup costs. The model is solved using a three-stage procedure that is embedded into a greedy heuristic. The first stage finds a feasible solution using the earliest-due-date (EDD) rule. The second stage then tries to improve this solution with respect to an objective of minimizing the total tardiness in filling the orders. An improvement procedure that tries to reduce the transport costs without increasing tardiness is used last. The algorithm was tested on real data from the company and on randomly generated instances. The largest instance contained 805 requests, 225 railcars, and 205 stations. The system yielded a significant cost reduction but computing times exceeded several hours in some experiments.

HOLMBERG, JOBORN, and LUNDGREN (1996) proposed a multicommodity network flow model for operational distribution of empty cars. Each commodity corresponds to a type of car, and linking constraints impose limits on the total number of empty cars that may be part of each scheduled train. Train movements are represented on a time–space network. The objective of the model is to minimize transportation and car shortage costs. The value of having a car in inventory at a given terminal after the planning period is also taken into account. A multiperiod planning horizon is considered and the operational model is solved using a sliding horizon framework in which decisions associated with the initial segment of the period are implemented whereas the others are reviewed by solving the model over the next segment. The model may also be used at the strategic level to evaluate the consequences of variations in the fleet size. A Lagrangian heuristic method was compared with a simple branch-and-bound procedure. Results obtained on real-life and randomly generated instances led to the conclusion that the model is very tractable. The largest instance solved contained 100 terminals and 20 car types. Substitution possibilities are also treated by extending the basic formulation, but no specific results were given for this extension. More details on this approach were given by JOBORN (1995) who also presented an analysis of empty freight car distribution at Swedish State Railways. An approach to determine train frequencies to minimize total costs for running trains and distributing empty cars was also introduced in a related paper by FLISBERG et al. (1996).

2.3.2 Multiple Railroad Models

A traditional repositioning strategy for freight cars consists of returning each unloaded car to its original loading point. This is a very simple and convenient approach given that a significant portion of freight shipments are made from the territory of one railroad to that of another. In the hope of reducing costs associated with empty movements, the concept of car pooling has gradually been introduced. Under a pooling agreement, railroads and shippers agree that cars unloaded at destination can be sent to any of a set of loading points.

A transshipment model to determine daily repositioning decisions that minimize network-wide costs was proposed by KIKUCHI (1985). GLICKMAN and SHERALI (1985) described two optimization approaches for the distribution of pooled cars that focus, respectively, on the benefits to the system as a whole and on the benefits to the individual railroads.

More recently, ADAMIDOU, KORNHAUSER, and KOSKOSIDIS (1993) argued that the problem of finding a global profit-maximizing distribution strategy for railroads sharing a fleet of cars is best represented as a generalized Nash equilibrium model. Their model includes coupling variables that link the individual multicommodity flow subproblems of the railroads and is solved through a Gauss–Seidel algorithm that iterates between these subproblems. When solving the subproblem for a particular railroad, the coupling variables are fixed using the optimal flows obtained when last solving the subproblems for all other railroads. The approach was tested on a large-scale, three-railroad instance generated from actual data, and appeared to be fast and robust. Different solution strategies were compared as well as various demand conditions.

The pooling of railcars used for the transportation of automobiles was studied by SHERALI and TUNCBILEK (1997) who proposed static and dynamic models for the fleet sizing problem. The static model tends to underestimate the real fleet size required because it is based on time-independent data. The dynamic model is based on a time–space network that represents the movement of empty cars between origins and destinations over the given planning horizon, with an objective of minimizing the fleet size required to satisfy all demands at different points in time. The problem is solved by decomposing the model into a series of smaller subproblems with a shorter, overlapping, temporal horizon. Once a subproblem is solved, the decisions for the initial part of the considered horizon are fixed, and the next subproblem is solved with the augmented flows. Test data instances generated randomly with realistic assumptions were used to evaluate the performance of the algorithm. The models have also been used successfully by the Association of American Railroads.

3. SCHEDULING PROBLEMS

WHEREAS THE MODELS of Section 2 are mainly concerned with the efficient routing of trains and freight, scheduling models address the temporal dimension of railroad operations. Because the physical rail network is shared by a large number of trains, it is indeed necessary to synchronize their use of the available resources. Also, the scheduling of freight and passenger train movements has an important impact on the quality and level of service provided. Finally, the scheduling of transportation activities is highly dependent upon the availability of rail equipment, such as the locomotives and passenger cars, that are needed to operate trains.

Compound models reviewed in Section 2.2.3 are an attempt at integrating the routing and scheduling aspects of rail freight transportation. However, these two closely intertwined problems are most often treated separately: operating plans are developed first, followed by train schedules that specify tentative departure and arrival times for the planned trains. The actual dispatching of trains is then performed by taking line capacity and other operational factors into account. This dispatching must often be performed simultaneously with the dispatching of passenger trains that operate in strict accordance with a timetable.

Most early models for train scheduling considered a set of stations connected by a single line. For example, the problem of developing timetables for passenger trains on a line of stations was studied by NEMHAUSER (1969) and SALZBORN (1969). The minimization of the number of railcars needed in a system of radial lines converging to a central station was also studied by SALZBORN (1970). Finally, an efficient approach for allocating demand to regular and express trains when delivering freight on a line network was suggested by ASSAD (1982).

More recently, the problem of finding a periodic train timetable that minimizes total passenger waiting time in stations of a network has received a lot of attention in the literature. Optimization models for that purpose were proposed by CEDER (1991), NACHTIGALL (1996), NACHTIGALL and VOGET (1996) and ODIJK (1996). The strategic problem of choosing a set of operating lines and their frequencies to serve demand and maximize the number of travelers on direct connections was studied by BUSSIECK, KREUZER, and ZIMMERMANN (1996). Also, ZWANEVELD et al. (1996) and KROON, ROMEIJN, and ZWANEVELD (1997) have proposed models and algorithms for the related problem of routing trains through railway stations. These contributions were reviewed in detail by BUSSIECK, WINTER, and ZIMMERMANN (1997), who discussed models for several discrete optimization problems in public rail transport. On a similar topic, NACHTIGALL (1995) discussed a problem that appears in passenger information systems and consists of computing shortest paths in a network with arc lengths that vary through time. Finally, NACHTIGALL and VOGET (1997) discussed a model for choosing the track segments to be upgraded to reduce train running times and thus minimize total passenger waiting time.

The following section contains a brief review of analytical models developed to measure the performance of a line relative to the traffic it accommodates, its configuration, operating policies, or other factors. Optimization models for train dispatching are then discussed, followed by models for locomotive assignment.

3.1 Analytical Line Models

Several models were proposed to estimate the delay to each train caused by interference on a rail line as a function of dispatching policies, traffic distribution, and physical track topology. Early results were given by FRANK (1966) for the case of a single-track line with two-way traffic but a single train speed and equally-spaced sidings. A more elaborate model was then developed by PETERSEN (1974) for trains of different speeds in each direction and sidings that allow for both meets and overtakes. His model assumes uniform and independent distributions of trains in each speed class over the considered horizon. The mean running times for trains in each class are obtained by solving a set of linear equations. Expressions for the expected meet and overtake interference delays on a partially double-tracked line were also developed by PETERSEN (1975). Necessary and sufficient conditions to guarantee that line blocking does not occur were given by PETERSEN and TAYLOR (1983). Queuing models to determine the expected dispatching delays on a single-track line with low-speed traffic and widely-spaced sidings were also described by GREENBERG, LEACHMAN, and WOLFF (1988). Finally, KRAFT (1988) extended Petersen's approach to take multiple train interactions into account and compared the results with myopic and optimized train dispatching.

CHEN and HARKER (1990) studied a more realistic problem in which trains have scheduled departure and arrival times instead of being randomly distributed over the planning horizon. The mean and variance of travel time are estimated by solving a system of nonlinear equations that also take into account uncertainties regarding actual departures. The extension of this framework to a partially double-tracked line was later presented by HARKER and HONG (1990).

Recently, HALLOWELL and HARKER (1996) described a model used to predict on-time arrival performance of trains on a partially double-tracked line with scheduled traffic. This model is an interesting alternative to simulation methods for estimating the lateness of delayed trains and can be used in tactical train scheduling or in train dispatching applications. In particular, it can be calibrated to generate target arrival times that can be achieved under an optimal planning of meets and overtakes.

The problem of track time use can also be seen from a game-theoretic standpoint. For example, HARKER and HONG (1994) presented an equilibrium model of an internal market for track time allocation. The generalized Nash equilibrium of the resulting model can be obtained by solving a quasivariational inequality problem.

Most line delay models assume a fixed track configuration. However, PETERSEN and TAYLOR (1987) presented a method for finding the optimal location and length of sidings for a single-track line with high-speed passenger trains. The solution is derived under the hypothesis of ideal train performance, but an analysis of robustness to small and large delays is also presented. Simulation experiments were performed using a methodology, introduced by PETERSEN and TAYLOR (1982), which is a framework for modeling train movements over single-track and multiple-track lines.

TABLE III
Characteristics of Train Dispatching Models

Authors	Problem Type	Planning Horizon	Objective Function	Model Structure	Solution Method
Jovanović and Harker (1991)	Fixed velocity	Tactical	Max reliability	MIP	Branch-and-bound
Carey and Lockwood (1995)	Fixed velocity	Operational	Min schedule deviation	Linear MIP	Heuristic decomposition
Carey (1994)	Fixed velocity	Operational	Min schedule deviation	Linear MIP	Heuristic decomposition
Carey (1994)	Fixed velocity	Operational	Min schedule deviation	Linear MIP	Heuristic decomposition
Kraay and Harker (1995)	Fixed velocity	Tactical	Min schedule deviation	Nonlinear MIP	Heuristic decomposition
Brännlund et al. (1996)	Fixed velocity	Tactical	Min schedule deviation	Linear IP	Lagrangian relaxation
Nõu (1997)	Fixed velocity	Tactical	Min schedule deviation	Linear IP	Lagrangian relaxation
Kraay et al. (1991)	Variable velocity	Tactical	Min train delays and fuel costs	Nonlinear MIP	Heuristic
Higgins et al. (1996)	Variable velocity	Operational	Min train delays and operating costs	Nonlinear MIP	Branch-and-bound
Higgins et al. (1997)	Variable velocity	Strategic	Min conflict delay and risk of delay	Nonlinear MIP	Heuristic decomposition

Finally, ÖZEKICI and ŞENGÖR (1994) analyzed the problem of train dispatching with the emphasis on suburban passenger rail transport systems. They considered a train station at which passenger arrivals, although random, are related to train departures through the published timetable (ÖZEKICI, 1987). The model is used for evaluating the performance, as measured by the service delay and the average waiting time of passengers, of different train dispatching strategies.

3.2 Train Dispatching Models

The train dispatching problem has received increased attention lately as several railroads are now developing and implementing advanced train control systems that provide real-time information on train position and velocity, as well as decisions to assist operations. These systems should help to reduce energy consumption and increase railroad line capacity and service reliability with improved train dispatching. An introduction to computerized train dispatching was written by PETERSEN, TAYLOR, and MARTLAND (1986). SMITH (1990) exposed the general guidelines that should be followed in designing a module for meet/pass planning. JOVANOVIĆ and HARKER (1990) also presented some analysis on the proper elaboration of computer-aided train dispatching systems. Then, HARKER (1989, 1995) reviewed some models and algorithms developed for such systems, and discussed the importance of advanced train control in the context of the current restructuring of technology and management practices that is taking place in the railroad industry.

Although most optimization models for train dispatching have appeared in the last decade, other enumerative approaches have also been in use. In particular, SZPIGEL (1973) described a method for train dispatching on a single-track line with meets and overtakes. SAUDER and WESTERMAN (1983) proposed a decision support system for train dispatching that implicitly enumerates all feasible meet locations and selects the one minimizing delays. KRAFT (1987) presented a branch-and-bound approach for resolving train conflicts to minimize a weighted sum of delays.

Computerized tools have also been developed to assist planners in constructing feasible dispatch plans. Such systems were described, for example, by RIVIER and TZIEROPOULOS (1984, 1987) and CHURCHOD and EMERY (1987).

We now discuss the recent optimization models for train dispatching. We first review models that assume all trains are operating at their maximum velocity, followed by models for the case in which velocity is variable. A summary of these models is presented in Table III.

3.2.1 Fixed Velocity Models

The aim of train dispatching models is to determine where trains will meet and pass to minimize train delays or deviations from the planned schedule while satisfying a set of operational constraints. Because the meeting and passing of trains is intimately related to their operating speed, a complete model should treat velocity as a decision variable. However, most dispatching models use a sequential approach and assume that trains will operate at maximum velocity whenever possible. A velocity profile is later determined for each train individually.

JOVANOVIĆ and HARKER (1991) proposed the SCAN system for the tactical scheduling of trains and maintenance operations. The main goal of their approach is to help in the design of reliable schedules in the sense that they are robust under stochastic operating conditions. The time horizon considered is a single day. The system, which can deal with single and double track segments, starts with a proposed schedule and first verifies its feasibility by separately analyzing each line of the network. To verify feasibility over a given line, a mixed integer programming problem with no explicit objective is solved with a branch-and-bound procedure to generate a feasible plan of meets and overtakes. This procedure incorporates a simulation method to model train movements and interactions. An automatic update procedure also helps in modifying an infeasible schedule into a feasible one. The mixed-integer programming problem has binary variables that indicate the ordering of the trains and continuous variables that represent departure and arrival times of trains at meetpoints. A complex set of constraints impose logical conditions concerning the meeting, passing, and following of trains. Time window constraints on the arrival and departure of each train are also present. The system performed well on a real-life network with 24 lines and schedules for 100 freight and passenger trains. The authors also report an implementation at a major U.S. railroad.

CAREY and LOCKWOOD (1995) described a model for the train dispatching problem on a line composed of several links connected by stations where overtaking can take place. The line is dedicated to traffic in one direction but trains operate at different speeds. Their model is a 0–1 mixed integer program that incorporates several headway constraints, bounds on departure and arrival times, and additional constraints used to strengthen the model. The *headway* is the time or distance separating two trains on the same link. The objective function to be minimized is rather general and takes deviations from the preferred schedule into account. The authors proposed to solve the model using a heuristic approach that first dispatches trains one at a time to obtain an initial solution, and then possibly redispatches individual trains to improve this solution. The subproblem of dispatching a single train has a reduced number of binary variables because the sequence order of the already dispatched trains is held fixed while the timings are allowed to vary. This problem is solved using a branch-and-bound procedure with branching decisions made on the link variables that specify the sequence order of the trains. Various strategies were proposed to accelerate the solution of the subproblem. In particular, branching in a depth-first search on the variables associated with the links in the same order as they are traversed by the trains seems to dramatically reduce the computing times. Of course, this method does not guarantee the optimality of the produced solution, nor does it ensure that a feasible solution will be found even if one exists. Good results are reported for computational experiments on small instances with 10 trains and 10 links.

In a follow-up paper, CAREY (1994a) extended the original model to introduce choices among multiple lines in each direction and choices of platforms to use for departures, arrivals, and stops at stations. This is done by introducing a more general type of link with two special cases representing train links and stations. Again, the model is solved with a heuristic decomposition approach that dispatches trains one at a time and redispatches individual trains until no further improvement in the solution is possible. Finally, the extension from one-way to two-way tracks was done by CAREY (1994b), who showed that the same solution methodology still applies in that case.

A model for optimizing freight train schedules was proposed by KRAAY and HARKER (1995). The goal of their approach is to provide a link between tactical train scheduling and actual operations by generating target times to be used in dispatching models such as the SCAN system (Jovanović and Harker, 1991). The model, which is a large nonlinear, mixed-integer program, directly considers the current position and relative importance of each train. Its solution indicates the target time for each train at each important point in its itinerary. For given values of the integer variables that determine the meeting and passing of trains, the model reduces to a continuous variable subproblem that is solved with an algorithm combining restricted simplicial decomposition and network flows. A simple heuristic approach and local search methods can be used to determine feasible values for the integer variables. Comparisons on a large set of real-life instances showed that the local search heuristics produced better results than the simple heuristic but required excessive computing time.

BRÄNNLUND et al. (1996) proposed a model to determine a profit maximizing schedule in which profit is measured by estimates of the value of running different types of services at specified times. The problem is formulated as a large integer programming problem and is solved with a Lagrangian relaxation approach in which track capacity constraints are dualized. The relaxed problem thus decomposes into a shortest path problem in a space–time network for each individual train. Feasible solutions are obtained with a heuristic that sequentially dispatches each train according to a priority list, given the current dual prices associated with track capacity constraints. Various dual optimization schemes were compared on instances with 26 and 30 trains on a single-track line connecting 17 stations. Computational experiments indicated that feasible solutions within a few percent of the lower bound were found in rather short computing times. According to this computational experience, the duality gap appears to increase as the line becomes more congested. Even though the approach is described for a single-track line, it easily extends to a double-tracked one.

In a follow-up paper, NÕU (1997) suggested and compared alternative approaches for generating feasible solutions. The author first extended the priority list heuristic described previously by Brännlund et al. In particular, a tabu search heuristic was proposed for the problem of finding the best possible permutation of the trains. Then, a conflict resolution heuristic which treats conflicts in order of occurrence was described. Finally, a greedy local improvement heuristic was introduced. This heuristic considers a feasible solution and tries to improve it by performing changes that maintain feasibility while improving the overall profit associated with the schedule. Computational experiments were performed on the same data that were used by Brännlund et al. Solution quality improvements in the order of 1% were obtained while computation times remained rather similar. The author concluded that the most effective approach is an enhanced priority list heuristic with a tabu search procedure to update the list.

3.2.2 *Variable Velocity Models*

Models that treat velocity as a decision variable are not very common even though they represent a significant improvement over fixed velocity models. Indeed, by treating operating speed endogenously, such models not only minimize deviations from the schedule but also quantify and minimize fuel consumption.

KRAAY, HARKER, and CHEN (1991) treated a train pacing problem in which train velocity and meeting and passing schedules are determined together to minimize fuel consumption and delays while satisfying time windows on the departure and arrival of each train. Their formulation is a nonlinear mixed integer program with a convex objective function. First, the authors proposed a branch-and-bound algorithm in which the initial relaxation is obtained by linearizing the objective function and by ignoring train interactions. This relaxation decomposes into simple linear programs solvable with a sorting routine and a line-search procedure. At each node of the branch-and-bound tree, cutting planes are added to gradually impose the relaxed constraints. When the relaxation solved at a node of the tree yields a feasible meet–pass plan, a feasible solution for the global problem can be computed by solving a nonlinear program in which the integer variables are held fixed. An alternative approach, based on the generation of feasible plans for the meeting and passing of trains, was also proposed. For each plan, the optimal velocity profiles are also computed by solving the nonlinear program with the integer variables being fixed. This approach is very convenient because it can use an oracle to generate plans that obey very complex constraints that do not even possess a mathematical representation. This approach can also be used to evaluate and rank different scenarios. Feasible meet–pass plans are generated using the logic of the SCAN system (Jovanović and Harker, 1991). Finally, the authors proposed a rounding heuristic to filter out meet–pass plans and retain only those closest to the optimal solution obtained when ignoring train interactions. Results on instances of a major railroad produced fuel savings in the order of 5% while the standard deviation in train arrival times decreased by more than 19%. A theoretical analysis shows that, as the number of sidings goes to infinity, the probability that the heuristic will give an optimal solution goes to one.

HIGGINS, KOZAN, and FERREIRA (1996) proposed a model and a solution method for the dispatching of trains on a single-track line. Their model mainly addresses the operational problem of dispatching trains in real time but can also serve at the strategic level to evaluate the impacts of timetable or infrastructure changes on train arrival times and train delays. The formulation is a complex nonlinear mixed integer program that incorporates lower and upper limits on train velocities for each train on each segment. The objective function seeks to minimize a combination of total train tardiness and fuel consumption. When a train will be delayed in a conflict at the next siding or has slack time, it will be paced to reduce fuel costs. The problem is solved using a branch-and-bound algorithm with lower bounds computed by using an estimate of the remaining delay cost, based on the calculation of the least cost path for each train. Comparisons with both an enumerative procedure that computes lower bounds by relaxing the remaining conflict constraints and a tabu search heuristic, showed that the proposed method is very effective at finding the optimal solution. A real-life instance with 31 trains and 14 sidings was solved in less than one minute. Other experiments are reported on instances of similar size.

In a follow-up paper, HIGGINS, KOZAN, and FERREIRA (1997) extended their solution methodology for simultaneously deciding the number and location of sidings and the optimal train schedule for a single-track line. This strategic problem is again

TABLE IV
Characteristics of Locomotive Assignment Models

Authors	Problem Type	Planning Level	Objective Function	Model Structure	Solution Method
Forbes et al. (1991)	Single engine	Tactical	Min operating costs	Assignment problems	Branch-and-bound
Fischetti and Toth (1997)	Single engine	Tactical	Min fleet size and deadheading	Assignment problems	Lagrangian relaxation
Florian et al. (1976)	Multiple engines	Strategic	Min investment and maintenance	Multicommodity	Benders decomposition
Smith and Sheffi (1988)	Multiple engines	Strategic	Min operating costs	Multicommodity	Heuristic
Chih et al. (1990)	Multiple engines	Operational	Max expected profit	Multicommodity	Heuristic decomposition
Ziarati et al. (1997)	Multiple engines	Operational	Min operating costs	Multicommodity	Dantzig–Wolfe decomposition
Nõu et al. (1997)	Multiple engines	Tactical	Min operating costs	Multicommodity	Dantzig–Wolfe decomposition
Ziarati et al. (1997)	Multiple engines	Operational	Min delays	Multicommodity	Dantzig–Wolfe decomposition
Cordeau et al. (1998)	Engines and cars	Tactical	Min operating costs	Multicommodity	Benders decomposition

modeled as a nonlinear mixed integer program. It is solved with a heuristic decomposition scheme that iterates between two subproblems until no further improvement is possible. The first subproblem chooses the positions of the sidings and the departure and arrival times for a given fixed schedule; the second subproblem chooses a train schedule, considering fixed siding locations. The method also considers an initial set of sidings that are held at fixed position. Because maximum train velocity on a given segment depends on the sidings location, velocity is determined endogenously. The objective function minimizes a weighted combination of conflict delay and risk of delay. The risk of delay represents the likely delay caused by unexpected events. Computational experiments on instances with up to 30 trains indicated that the algorithm converges very quickly.

3.3 Locomotive Assignment Models

Given a planned train schedule, the locomotive assignment problem consists of assigning a set of locomotives to the scheduled trains to satisfy requirements expressed as a number of locomotives or as a measure of the pulling power needed (i.e., horsepower and tonnage). At a strategic planning level, the objective followed is usually to minimize the required fleet size. At the tactical and operational levels, the available rolling stock is given and one usually wants to minimize costs incurred by light running. Light running or *deadheading* occurs when an engine must be repositioned between two successive trips.

Early research on the problem of assigning engines to trains was conducted by CHARNES and MILLER (1957), who used linear programming for the assignment of crew–engine pairings to a set of potential trips to provide each train in a given schedule with sufficient resources. BARTLETT (1957) gave an algorithm for minimizing fleet size based on the idea that, for a fixed time horizon, this objective is tantamount to minimizing total idle time. An algorithm for finding an assignment that satisfies maintenance constraints while minimizing deviations from a target mileage between successive maintenance stops was proposed in related work by BARTLETT and CHARNES (1957).

Over the years, many railways have developed decision support systems to assist planners in making locomotive assignment and scheduling decisions. Although early systems relied in large part on simulation techniques and decision rules dictated by experience, some of them also used optimization methods. For example, GOHRING (1971) and MCGAUGHEY, GOHRING, and MCBRAYER (1973) described a periodic network flow model, solved with the out-of-kilter algorithm (Ford and Fulkerson, 1962), to minimize fleet size at Southern Railway. Also, HOLT (1973) mentioned the use of branch-and-bound procedures and decomposition approaches for locomotive distribution at British Railways.

We now review the more recent optimization models for locomotive assignment. We first discuss the case in which each train needs a single engine, followed by models for the multiple engine case. The simultaneous assignment of both engines and cars to passenger trains is treated last. Table IV provides a summary of the reviewed models.

3.3.1 Single Locomotive Models

Most models for the problem in which multiple engine types are available but each train needs a single locomotive have a multicommodity network flow structure with linking constraints that ensure that each train is covered exactly once. For example, BOOLER (1980) proposed a heuristic algorithm that starts with a feasible allocation of locomotive types to the trains and iteratively updates this allocation using the dual information gathered when solving the resulting assignment problems. A Lagrangian relaxation approach, that dualizes the linking constraints in the objective function, was later proposed by the same author (BOOLER, 1995). WRIGHT (1989) compared stochastic algorithms based on the solution of assignment problems and the update of the locomotive types assigned to the trains.

An exact algorithm for a model with a similar structure was proposed by FORBES, HOLT, and WATTS (1991). The objective function takes into consideration fixed costs and operating costs. The solution technique consists of solving the LP relaxation of an integer programming formulation before applying a branch-and-bound procedure to obtain an integer solution. To solve the continuous relaxation, a further relaxation is obtained by removing the locomotive type restrictions. The solution to that problem is then converted into a dual feasible solution to the original problem and the dual simplex method is used to obtain the optimal solution to the LP relaxation. Branching is first performed on the number of locomotives used. Additional branching is performed on the successors of the trains and on the locomotive types assigned to the trains. The datasets used for testing purposes did not impose constraints on the number of available locomotives of each type, but the authors mentioned how these can be enforced in their formulation. They reported very small integrality gaps, in particular when the objective function does not include preferences for locomotive types.

Very recently, a heuristic method for the weekly problem was proposed by FISCHETTI and TOTH (1997). Engines are distributed across a number of depots that are associated with stations of the network. Each depot has a maximum number of engines available and each engine must go through its depot every week to allow for maintenance. In addition, engine trips must satisfy a set of operational constraints. By relaxing the maintenance and operational constraints, one obtains an assignment problem whose solution provides a very good lower bound on the optimal solution. The objective function is a weighted combination of the number of engines needed, the number of deadheading trips performed, and the distance covered by deadheading trips. Real-life instances with up to 10,000 trains were solved in less than one hour on a workstation computer. Cost savings in the order of 10–20% were typically obtained over the solution in use by the Italian Railways.

3.3.2 Multiple Locomotive Models

When each train may require more than one locomotive but these requirements are given as a number of engines, the problem can still be formulated as a multicommodity network flow problem with rather simple linking constraints. The most difficult version of the problem occurs when multiple locomotive types are available and each train may require more than one locomotive to satisfy its requirements expressed in terms of motive power.

One of the first models dealing with this version of the problem was proposed by FLORIAN et al. (1976). The strategic problem considered is to select the mix of engine types that gives the lowest capital investment and maintenance costs over a long planning horizon, while providing each train with sufficient engines to meet its motive power requirements. In this model, the motive power requirements of each train are determined according to its weight and length in terms of cars, and to the route on which it must travel. The model used is defined on a set of network flow circulation problems with some linking constraints that translate the motive power requirements. The solution approach is based on Benders decomposition (BENDERS, 1962) and takes advantage of the particular structure of the problem. Variables in the integer programming master problem impose lower bounds on the arcs of network flow subproblems. To speed up the solution of the master problem, a decomposition scheme is coupled with a rounding heuristic. Upper bounds on the number of engines of each type are not treated. Computational results were reported on problems with a few hundred trains and the convergence was deemed slow on the larger instances. However, it should be emphasized that the algorithm was stopped after less than 30 iterations were performed. Hence, given the performance of today's computers, it is very likely that the conclusions would now be different.

A model that incorporates the uncertainty in locomotive requirements was suggested by SMITH and SHEFFI (1988). This model has a multicommodity network flow structure with linking constraints that enforce locomotive requirements expressed as a lower bound on the horsepowers supplied to each train. These constraints are relaxed in the objective function by using a penalty function that permits deviations from the requirements at a cost. Also, the lower bounds on horsepower are replaced by random variables with known distributions. The resulting model has a convex nonlinear cost function and is solved with a two-phase heuristic. In the first phase, a feasible solution is obtained by incremental flow assignments along shortest paths. In the second phase, interchanges are performed to improve the solution by identifying cycles with a negative marginal cost. The major advantage of this heuristic procedure is that it maintains integrality throughout. To evaluate its performance, lower bounds were computed with two approaches. The first one relaxes integrality constraints and solves the resulting problem with a Frank–Wolfe method (Frank and Wolfe, 1956). The second one uses a piecewise linearization of the cost function to obtain a pure network flow problem. Computational experiments on instances with up to 102 trains produced feasible solutions with short computing times and costs within a few percent of the best lower bound.

CHIH et al. (1990) described the implementation of an operational planning model for locomotive assignment. The model, which seeks to maximize the difference between expected revenue and operational costs, is based on a time–space network representing all possible locomotive movements during the planning horizon. To obtain a first approximation of motive power assignments to a set of weekly scheduled trains, a multicommodity network flow problem is solved with a resource-directive decomposition approach. Locomotives that must be directed to a shop for maintenance are then routed individually by solving a shortest path problem and horsepower requirements are lowered to reflect these assignments. Given the solution to the multicommodity network flow problem and the residual requirements, locomotive consists are finally built for each train by an exhaustive enumeration process. The approach was tested on actual data from the Union Pacific Railroad. On an instance with 15 types of locomotives and a network for each type containing more than 25,000 arcs, a solution was found within 30 minutes.

More recently, a Dantzig–Wolfe decomposition approach (Dantzig and Wolfe, 1960) was developed for the operational version of this problem by ZIARATI et al. (1997). Train requirements are determined as above but some engines must also be dispatched to special stations at which they have to perform local work. A list of preferred locomotives is considered for each train and care is also taken of the locomotives that must be routed to a shop for maintenance. The objective considered is the minimization of the total operational costs. The problem is modeled as a multicommodity flow problem with supplementary variables and constraints. The time horizon considered is a week, but, to solve very large instances, the problem is divided on a temporal basis into a set of overlapping slices involving fewer trains. Once the problem for a slice is solved, the problem for the next one is solved with initial conditions determined by the solution of the preceding slice. The problem for each slice is solved using a branch-and-bound procedure in which the linear relaxations are solved by column generation. Constrained and unconstrained shortest path problems must be solved on an acyclic network to generate columns for the master problem. A heuristic branching strategy is used, in which many path variables are fixed together. Branching decisions are made on the path variables with the largest fractional part, and the selected variables are rounded up to the next integer. Computational experiments carried out on real-life data involving approximately 2000 trains allowed an improvement of 7% over the solution in use by the company when taking slices of two days with a one-day overlap. This improvement goes to 7.5% when slices of three days are used, but the computations then take a few hours. ZIARATI et al. (1998) introduced additional cuts, based on the enumeration of feasible assignments of locomotive combinations to trains, which strengthen the LP relaxation lower bound and improve solution quality. A day-to-day operational model was also proposed by ZIARATI (1997).

A similar approach was used by NÕU, DESROSIERS, and SOUMIS (1997) for the tactical assignment problem at Swedish State Railways. In this problem, cyclic locomotive assignments are sought and maintenance constraints related to cumulated distance must be satisfied. Two approaches based on a branch-and-bound procedure and Dantzig–Wolfe decomposition are presented for solving the problem. In the first approach, the weekly problem is replaced by a series of smaller size problems with overlapping horizons. In the second one, maintenance constraints are relaxed to obtain a smaller problem that is solved without being decomposed on a temporal basis. Tests performed with actual data from Swedish State Railways involving 2422 trains showed that the first approach failed to produce a feasible cyclic solution. The second approach produced a solution that violated maintenance constraints for a very limited number of locomotives.

In some cases, the operational locomotive assignment problem may be infeasible because not enough engines are available. One possibility to circumvent this difficulty is to allow train undercovering. Undercovering happens when the motive power requirements are not fully satisfied. This is easily achieved by introducing slack variables in the appropriate constraints. ZIARATI, SOUMIS, and DESROSIERS (1997) presented an alternative approach that consists of delaying trains. The basic idea of the method is to postpone the departure of an undercovered train until enough locomotives are available at the origin station. For the case of express trains, one can instead postpone the departure of a preceding train to assign the available engines to the express train. Using these strategies, one can often find a feasible solution in terms of the covering constraints. To determine a valid lower bound, the authors use an augmented network that includes both penalty and delay costs, as well as fixed and routing costs. The solution to an instance with almost 2000 train segments was obtained in less than 10 minutes and had a cost within 5% of the lower bound.

3.3.3 Locomotive and Passenger Car Models

Very little work has been accomplished concerning the assignment of locomotives and cars in the context of passenger transportation. A decision support system was developed by RAMANI and MANDAL (1992) for the planning of passenger trains at Indian Railways. However, the assignment of locomotives and cars is dealt with separately and the system uses a simple local improvement procedure that generates optimal train connections by examining the departures and arrivals at individual stations. This procedure is reminiscent of the algorithm given by Bartlett (1957) for fleet size minimization. The work of Ramani and Mandal extends an information system for car assignment developed by RAMANI (1981).

Recently, an optimization model for the assignment of both locomotives and passenger cars was proposed by CORDEAU, SOUMIS, and DESROSIERS (1998). The tactical periodic problem is formulated as an integer programming problem based on a time–space network. As in the work of Florian et al. (1976) for locomotive assignment, this model possesses an interesting variable decomposition: for given values of the binary variables that represent the assignment of equipment combinations to trains, the problem decomposes into one network flow problem for each type of equipment. The formulation incorporates compatibility constraints between the different types of equipment that may be combined to form valid train consists. Equipment availability for each type is also enforced. However, the model does not directly impose maintenance constraints. Comparisons between primal and dual decomposition methods and a simplex-based branch-and-bound approach showed that the formulation was best solved using Benders decomposition (Benders, 1962). Algorithmic refinements were suggested to improve the performance of the algorithm. In computational experiments performed on real-life data, the algorithm found optimal solutions within short computation times. The largest instance solved contained six types of equipment and 348 trains over a period of one week.

4. CONCLUSIONS

THIS PAPER HAS presented a review of the recent optimization models proposed for solving routing and scheduling problems in rail transportation. The field is clearly receiving increased attention as measured by the number of contributions in the last few years. The nature and scope of the research conducted is also gaining in diversity as nearly every domain of rail transport planning has been the object of some recent research.

There also appears to be a constant refinement and diversification of the modeling and solution methods proposed and used. Early models were usually built to have a structure that made them solvable by linear programming or network optimization. One then witnessed a gradual introduction of integer programs with simple underlying structures. Whereas some recent models are solved with more sophisticated mathematical programming techniques, others still are solved using meta-heuristics that have proven to be very effective for several classes of discrete optimization problems. Of course, this progression is also made possible by the increased power of computers and information systems.

As mentioned in the introduction, optimization models for train routing and scheduling have advanced tremendously in the last few years. Although early models were often based on very crude approximations of reality, recent applications demonstrate an important effort to deal with complex yet important characteristics of the actual functioning of railway systems. As a result, problems which, in the past, were only approachable by simulation can now be solved, at least approximately, using mathematical optimization. Nevertheless, simulation techniques have also made considerable progress in the last decade and remain a very useful tool of analysis and support to decision making. The recent work of POWELL (1995) is an illustrative example of this progress.

Also, despite the increasing realism of optimization models, considerable work remains to be accomplished to make the railways benefit from this wealth of knowledge. Even though most proposed models are tested on realistic data instances, very few are actually implemented and used in railway operations. Hence, efforts must be made to bridge the gap between theory and practice. MARTLAND and SUSSMAN (1995) presented an interesting discussion of factors that explain the success or failure of different approaches.

Future research paths in rail transportation planning are oriented toward models that address the integration of various policies. Because rail activities are generally complex and involve large-scale systems, the traditional approach in the industry has been to separate planning activities into several components. This natural tendency yields more manageable subsystems but also presents several limitations. In particular, there is a strong incentive to simultaneously treat routing and scheduling problems because of the important interactions linking these two categories of decisions. Hence, models that integrate several aspects and levels of planning should be increasingly common in upcoming years.

ACKNOWLEDGMENTS

WE SINCERELY THANK Dr. Edwin R. Kraft of Amtrak. His suggestions helped us to improve both the contents and presentation of this paper. Thanks are also due to Professor Jacques Desrosiers for his comments on preliminary versions of the text. This work was supported by the Québec Government (Fonds pour la Formation de Chercheurs et l'Aide à la Recherche) and by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

E. A. ADAMIDOU, A. L. KORNHAUSER, AND Y. A. KOSKOSIDIS, "A Game Theoretic/Network Equilibrium Solution Approach for the Railroad Freight Car Management Problem," *Transp. Res.* **27B,** 237–252 (1993).

W. P. ALLMAN, "An Optimization Approach to Freight Car Allocation Under Time-Mileage Per Diem Rental Rates," *Management Sci.* **18B,** 567–574 (1972).

A. A. ASSAD, "Modelling of Rail Networks: Toward a Routing/Makeup Model," *Transp. Res.* **14B,** 101–114 (1980a).

A. A. ASSAD, "Models for Rail Transportation," *Transp. Res.* **14A,** 205–220 (1980b).

A. A. ASSAD, "Analytical Models in Rail Transportation: An Annotated Bibliography," *INFOR* **19,** 59–80 (1981).

A. A. ASSAD, "A Class of Train-Scheduling Problems," *Transp. Sci.* **16,** 281–310 (1982).

A. A. ASSAD, "Analysis of Rail Classification Policies," *INFOR* **21,** 293–314 (1983).

B. AVI-ITZHAK, B. A. BENN, AND B. A. POWELL, "Car Pool Systems in Railroad Transportation: Mathematical Models," *Management Sci.* **13,** 694–711 (1967).

Z. Ž. AVRAMOVIĆ, "Method for Evaluating the Strength of Retarding Steps on a Marshalling Yard Hump," *Eur. J. Operational Res.* **85,** 504–514 (1995).

T. E. BARTLETT, "An Algorithm for the Minimum Number of Transport Units to Maintain a Fixed Schedule," *Naval Res. Logist. Quart.* **4,** 139–149 (1957).

T. E. BARTLETT AND A. CHARNES, "Cyclic Scheduling and Combinatorial Topology: Assignment and Routing of Motive Power to Meet Scheduling and Maintenance Requirements. Part II. Generalization and Analysis," *Naval Res. Logist. Quart.* **4,** 207–220 (1957).

G. J. BEAUJON AND M. A. TURNQUIST, "A Model for Fleet Sizing and Vehicle Allocation," *Transp. Sci.* **25,** 19–45 (1991).

M. BECKMANN, C. B. MCGUIRE, AND C. B. WINSTON, *Studies in the Economics of Transportation,* Yale University Press, New Haven, CT, 1956.

J. F. BENDERS, "Partitioning Procedures for Solving Mixed-Variables Programming Problems," *Numer. Math.* **4,** 238–252 (1962).

L. D. BODIN, B. L. GOLDEN, A. D. SCHUSTER, AND W. ROMIG, "A Model for the Blocking of Trains," *Transp. Res.* **14B,** 115–120 (1980).

J. M. P. BOOLER, "The Solution of a Railway Locomotive Scheduling Problem," *J. Operational Res. Soc.* **31,** 943–948 (1980).

J. M. P. BOOLER, "A Note on the Use of Lagrangean Relaxation in Railway Scheduling," *J. Operational Res. Soc.* **46,** 123–127 (1995).

U. BRÄNNLUND, P. O. LINDBERG, A. NÕU, AND J. E. NILSSON, "Railway Timetabling using Lagrangian Relaxation," *Transp. Sci.* **32,** 000–000.

M. R. BUSSIECK, P. KREUZER, AND U. T. ZIMMERMANN, "Optimal Lines for Railway Systems," *Eur. J. Operational Res.* **96,** 54–63 (1996).

M. R. BUSSIECK, T. WINTER, AND U. T. ZIMMERMANN, "Discrete Optimization in Public Rail Transport," *Math. Programming* **79,** 415–444 (1997).

K. C. CAMPBELL, Booking and Revenue Management for Rail Intermodal Services, PhD dissertation, Department of Systems Engineering, University of Pennsylvania, Philadelphia, PA, 1996.

A. CAPRARA, M. FISCHETTI, P. TOTH, D. VIGO, AND P. L. GUIDA, "Algorithms for Railway Crew Management," *Math. Programming* **79,** 125–141 (1997).

M. CAREY, "A Model and Strategy for Train Pathing with Choice of Lines, Platforms and Routes," *Transp. Res.* **28B,** 333–353 (1994a).

M. CAREY, "Extending a Train Pathing Model from One-Way to Two-Way Track," *Transp. Res.* **28B,** 395–400 (1994b).

M. CAREY AND D. LOCKWOOD, "A Model, Algorithms and Strategy for Train Pathing," *J. Operational Res. Soc.* **46,** 988–1005 (1995).

A. CEDER, "A Procedure to Adjust Transit Trip Departure Times through Minimizing the Maximum Headway," *Comput. Operations Res.* **18,** 417–431 (1991).

A. CHARNES AND M. H. MILLER, "A Model for the Optimal Programming of Railway Freight Train Movements," *Management Sci.* **3,** 74–92 (1957).

B. CHEN AND P. T. HARKER, "Two Moments Estimation of the Delay on Single-Track Rail Lines with Scheduled Traffic," *Transp. Sci.* **24,** 261–275 (1990).

K. C. CHIH, M. A. HORNUNG, M. S. ROTHENBERG, AND A. L. KORNHAUSER, "Implementation of a Real Time Locomotive Distribution System," In *Computer Applications in Railway Planning and Management,* T. K. S. Murthy, R. E. Rivier, G. F. List and J. Mikolaj (eds.), Computational Mechanics Publications, Southampton, U.K., 39–49, 1990.

A. CHURCHOD AND D. EMERY, "Computer-Aided Planning for Major Railway Stations," In *Computers in Railway Installations, Track and Signalling*, T. K. S. Murthy, F. E. Young, S. Lehmann and W. R. Smith (eds.), Computational Mechanics Publications, Berlin, 3–19, 1987.

L. COOPER AND L. J. LEBLANC, "Stochastic Transportation Problems and Other Network Related Convex Problems," *Naval Res. Logist. Quart.* **24,** 327–337 (1977).

J.-F. CORDEAU, F. SOUMIS, AND J. DESROSIERS, A Benders Decomposition Approach for the Locomotive and Car Assignment Problem, Technical report G-98-35, GERAD, École des Hautes Études Commerciales de Montréal, Canada, 1998.

T. G. CRAINIC, "A Comparison of Two Methods for Tactical Planning in Rail Freight Transportation," In *Operational Research '84*, J. P. Brans (ed.), Elsevier Science Publishers, Amsterdam, 707–720, 1984.

T. G. CRAINIC, J.-A. FERLAND, AND J.-M. ROUSSEAU, "A Tactical Planning Model for Rail Freight Transportation," *Transp. Sci.* **18,** 165–184 (1984).

T. G. CRAINIC, M. FLORIAN, AND J.-E. LÉAL, "A Model for the Strategic Planning of National Freight Transportation by Rail," *Transp. Sci.* **24,** 1–24 (1990a).

T. G. CRAINIC, M. GENDREAU, AND P. DEJAX, "Modelling the Container Fleet Management Problem using a Stochastic Dynamic Approach," In *Operational Research '90*, H. E. Bradley (ed.), Pergamon Press, New York, 473–486, 1990b.

T. G. CRAINIC, M. GENDREAU, AND P. DEJAX, "Dynamic and Stochastic Models for the Allocation of Empty Containers," *Operations Res.* **41,** 102–126 (1993).

T. G. CRAINIC AND J.-M. ROUSSEAU, "Multicommodity, Multimode Freight Transportation: A General Modeling and Algorithmic Framework for the Service Network Design Problem," *Transp. Res.* **20B,** 225–242 (1986).

R. R. CRANE, F. B. BROWN, AND R. O. BLANCHARD, "An Analysis of a Railroad Classification Yard," *Operations Res.* **3,** 262–271 (1955).

C. F. DAGANZO, "Static Blocking at Railyards: Sorting Implications and Track Requirements," *Transp. Sci.* **20,** 189–199 (1986).

C. F. DAGANZO, "Dynamic Blocking for Railyards. Part I. Homogeneous Traffic," *Transp. Res.* **21B,** 1–27 (1987a).

C. F. DAGANZO, "Dynamic Blocking for Railyards. Part II. Heterogeneous Traffic," *Transp. Res.* **21B,** 29–40 (1987b).

C. F. DAGANZO, R. G. DOWLING, AND R. W. HALL, "Railroad Classification Yard Throughput: The Case of Multistage Triangular Sorting," *Transp. Res.* **17A,** 95–106 (1983).

G. B. DANTZIG AND P. WOLFE, "Decomposition Principle for Linear Programming," *Operations Res.* **8,** 101–111 (1960).

P. J. DEJAX AND T. G. CRAINIC, "A Review of Empty Flows and Fleet Management Models in Freight Transportation," *Transp. Sci.* **21,** 227–247 (1987).

J. M. FARVOLDEN AND W. B. POWELL, "Subgradient Methods for the Service Network Design Problem," *Transp. Sci.* **28,** 256–272 (1994).

M. FISCHETTI AND P. TOTH, A Package for Locomotive Scheduling, Technical report DEIS-OR-97-16, University of Bologna, Italy, 1997.

P. FLISBERG, K. HOLMBERG, M. JOBORN, AND J. T. LUNDGREN, A Model for Dimensioning of Transportation Capacities in a Railway Network," Technical report LiTH-MAT-R-1996-25, Department of Mathematics, Linköping University, Sweden, 1996.

M. FLORIAN, G. BUSHELL, J. FERLAND, G. GUÉRIN, AND L. NASTANSKY, "The Engine Scheduling Problem in a Railway Network." *INFOR* **14,** 121–138 (1976).

M. A. FORBES, J. N. HOLT, AND A. M. WATTS, "Exact Solution of Locomotive Scheduling Problems," *J. Operational Res. Soc.* **42,** 825–831 (1991).

L. R. FORD AND D. R. FULKERSON, *Flows in Networks.* Princeton University Press, Princeton, NJ, 1962.

M. FRANK AND P. WOLFE, "An Algorithm for Quadratic Programming," *Naval Res. Logist. Quart.* **3,** 95–110 (1956).

O. FRANK, "Two-Way Traffic on a Single Line of Railway," *Operations Res.* **14,** 801–811 (1966).

T. S. GLICKMAN AND H. D. SHERALI, "Large-Scale Network Distribution of Pooled Empty Freight Cars over Time, with Limited Substitution and Equitable Benefits," *Transp. Res.* **19B,** 85–94 (1985).

K. W. GOHRING, "Application of Network Flow Theory to the Distribution of Locomotives and Cabooses," Presented at the Institute of Management Sciences Southeastern Chapter Winter Symposium, March 3–4, 1971.

M. F. GORMAN, "An Application of Genetic and Tabu Searches to the Freight Railroad Operating Plan Problem," *Ann. Operations Res.* **78,** 51–69 (1998).

B. S. GREENBERG, R. C. LEACHMAN, AND R. W. WOLFF, "Predicting Dispatching Delays on a Low Speed, Single Track Railroad," *Transp. Sci.* **22,** 31–38 (1988).

A. E. HAGHANI, "Rail Freight Transportation: A Review of Recent Optimization Models for Train Routing and Empty Car Distribution," *J. Advanced Transp.* **21,** 147–172 (1987).

A. E. HAGHANI, "Formulation and Solution of a Combined Train Routing and Makeup, and Empty Car Distribution Model," *Transp. Res.* **23B,** 433–452 (1989).

S. F. HALLOWELL AND P. T. HARKER, "Predicting On-Time Line-Haul Performance in Scheduled Railroad Operations," *Transp. Sci.* **30,** 364–378 (1996).

P. T. HARKER, "Use of Advanced Train Control Systems in Scheduling and Operating Railroads: Models, Algorithms, and Applications," *Transp. Res. Rec.* **1263,** 101–110 (1989).

P. T. HARKER, "Services and Technology: Reengineering the Railroads," *Interfaces* **25,** 72–80 (1995).

P. T. HARKER AND S. HONG, "Two Moments Estimation of the Delay on a Partially Double-Track Rail Line with Scheduled Traffic," *J. Transp. Res. Forum* **31,** 38–49 (1990).

P. T. HARKER AND S. HONG, "Pricing of Track Time in Railroad Operations: An Internal Market Approach," *Transp. Res.* **28B,** 197–212 (1994).

H. HERREN, "The Distribution of Empty Wagons by Means of Computer: An Analytical Model of the Swiss Federal Railways," *Rail Internat.* **4,** 1005–1010 (1973).

H. HERREN, "Computer-Controlled Empty Wagon Distribution on the SBB," *Rail Internat.* **8,** 25–32 (1977).

A. HIGGINS, E. KOZAN, AND L. FERREIRA, "Optimal Scheduling of Trains on a Single Line Track," *Transp. Res.* **30B,** 147–161 (1996).

A. HIGGINS, E. KOZAN, AND L. FERREIRA, "Modelling the Number and Location of Sidings on a Single Line Railway," *Comput. Operations Res.* **3,** 209–220 (1997).

K. HOLMBERG, M. JOBORN, AND J. T. LUNDGREN, "Improved Empty Freight Car Distribution," *Transp. Sci.* **32,** 163–173 (1998).

J. HOLT, "Locomotive Scheduling by Computer Bashpeak, *Rail Internat.* **4,** 1053–1058 (1973).

C. L. HUNTLEY, D. E. BROWN, D. E. SAPPINGTON, AND B. P. MARKOWICZ, "Freight Routing and Scheduling at CSX Transportation," *Interfaces* **25(3),** 58–71 (1995).

M. JOBORN, Empty Freight Car Distribution at Swedish Railways—Analysis and Optimization Modeling, PhD dissertation, Department of Mathematics, Linköping University, Sweden, 1995.

W. C. JORDAN AND M. A. TURNQUIST, "A Stochastic, Dynamic Network Model for Railroad Car Distribution," *Transp. Sci.* **17,** 123–145 (1983).

D. JOVANOVIĆ AND P. T. HARKER, "A Decision Support System for Train Dispatching: An Optimization-Based Methodology." *J. Transp. Res. Forum* **31,** 25–37 (1990).

D. JOVANOVIĆ AND P. T. HARKER, "Tactical Scheduling of Rail Operations: The SCAN I System," *Transp. Sci.* **25,** 46–64 (1991).

M. H. KEATON, "Designing Optimal Railroad Operating Plans: Lagrangian Relaxation and Heuristic Approaches," *Transp. Res.* **23B,** 415–431 (1989).

M. H. KEATON, "Service-Cost Tradeoffs for Carload Freight Traffic in the U.S. Rail Industry," *Transp. Res.* **25A,** 363–374 (1991).

M. H. KEATON, "Designing Railroad Operating Plans: A Dual Adjustment Method for Implementing Lagrangian Relaxation," *Transp. Sci.* **26,** 263–279 (1992).

S. KIKUCHI, "Empty Freight Car Dispatching Model under Freight Car Pool Concept," *Transp. Res.* **19B,** 169–185 (1985).

D. R. KRAAY AND P. T. HARKER, "Real-Time Scheduling of Freight Railroads," *Transp. Res.* **29B,** 213–229 (1995).

D. R. KRAAY, P. T. HARKER, AND B. CHEN, "Optimal Pacing of Trains in Freight Railroads: Model Formulation and Solution," *Operations Res.* **39,** 82–99 (1991).

E. R. KRAFT, "A Branch and Bound Procedure for Optimal Train Dispatching," *J. Transp. Res. Forum* **28,** 263–276 (1987).

E. R. KRAFT, "Analytical Models for Rail Line Capacity Analysis," *J. Transp. Res. Forum* **29,** 153–162 (1988).

E. R. KRAFT, A Reservations-Based Railway Network Operations Management System, PhD dissertation, Department of Systems Engineering, University of Pennsylvania, Philadelphia, PA, 1998.

L. G. KROON, H. E. ROMEIJN, AND P. J. ZWANEVELD, "Routing Trains through Railway Stations: Complexity Issues," *Eur. J. Operational Res.* **98,** 485–498 (1997).

L. J. LEBLANC, "Global Solutions for a Nonconvex Nonconcave Rail Network Model," *Management Sci.* **23,** 131–139 (1976).

P. MAHEY, "Méthodes de Décomposition et Décentralisation en Programmation linéaire," *RAIRO Recherche opérationnelle* **20,** 287–306 (1986).

E. MANSFIELD AND H. H. WEIN, "A Model for the Location of a Railroad Classification Yard," *Management Sci.* **4,** 292–313 (1958).

A. MARÍN AND J. SALMERÓN, "Tactical Design of Rail Freight Networks. Part I. Exact and Heuristic Methods," *Eur. J. Operational Res.* **90,** 26–44 (1996a).

A. MARÍN AND J. SALMERÓN, "Tactical Design of Rail Freight Networks. Part II. Local Search Methods with Statistical Analysis," *Eur. J. Operational Res.* **94,** 43–53 (1996b).

B. P. MARKOWICZ AND M. A. TURNQUIST, "Applying the LP Solution to the Daily Distribution of Freight Cars," Presented at the TIMS/ORSA National Meeting, Las Vegas, 1990.

D. R. MARTINELLI AND H. TENG, "Optimization of Railway Operations using Neural Networks," *Transp. Res.* **4C,** 33–49 (1996).

C. D. MARTLAND, "PMAKE Analysis: Predicting Rail Yard Time Distributions using Probabilistic Train Connection Standards," *Transp. Sci.* **16,** 476–506 (1982).

C. D. MARTLAND AND J. M. SUSSMAN, "A Perspective on Rail Systems Modelling: What Works and What Doesn't," Working paper 95-1, MIT Center for Transportation Studies, Cambridge, MA, 1995.

R. S. MCGAUGHEY, K. W. GOHRING, AND R. N. MCBRAYER, "Planning Locomotive and Caboose Distribution," *Rail Internat.* **4,** 1213–1218 (1973).

V. B. MENDIRATTA AND M. A. TURNQUIST, "A Model for Management of Empty Freight Cars," *Transp. Res. Rec.* **838,** 50–55 (1982).

S. C. MISRA, "Linear Programming of Empty Wagon Disposition," *Rail Internat.* **3,** 151–158 (1972).

M.-H. MORIN, Modélisation et Décomposition des problèmes de transbordement dynamiques: Application à la répartition des Wagons Fret/SNCF, PhD dissertation, Université J. Fourier, Grenoble, France, 1993.

E. K. MORLOK AND R. B. PETERSON, "Final Report on a Development of a Geographic Transportation Network Generation and Evaluation Model," *J. Transp. Res. Forum* **11,** 71–105 (1970).

K. NACHTIGALL, "Time Depending Shortest-Path Problems with Applications to Railway Networks," *Eur. J. Operational Res.* **83,** 154–166 (1995).

K. NACHTIGALL, "Periodic Network Optimization with Different Arc Frequencies," *Discrete Appl. Math.* **69,** 1–17 (1996).

K. NACHTIGALL AND S. VOGET, "A Genetic Algorithm Approach to Periodic Railway Synchronization," *Comput. Operations Res.* **23,** 453–463 (1996).

K. NACHTIGALL AND S. VOGET, “Minimizing Waiting Times in Integrated Fixed Interval Timetables by Upgrading Railway Tracks,” *Eur. J. Operational Res.* **103,** 610–627 (1997).

G. L. NEMHAUSER, “Scheduling Local and Express Trains,” *Transp. Sci.* **3,** 164–175 (1969).

H. N. NEWTON, Network Design Under Budget Constraints with Application to the Railroad Blocking Problem, PhD dissertation, Auburn University, Auburn, AL, 1996.

A. NÕU, Railway Timetabling—Lagrangian Heuristics, Technical report TRITA/MAT-97-OS12, Royal Institute of Technology, Stockholm, Sweden, 1997.

A. NÕU, J. DESROSIERS, AND F. SOUMIS, Weekly Locomotive Scheduling at Swedish State Railways, Technical report G-97-35, GERAD, École des Hautes Études Commerciales de Montréal, Canada, 1997.

L. K. NOZICK AND E. K. MORLOK, “A Model for Medium-Term Operations Planning in an Intermodal Rail-Truck Service,” *Transp. Res.* **31A,** 91–107 (1997).

M. A. ODIJK, “A Constraint Generation Algorithm for the Construction of Periodic Railway Timetables,” *Transp. Res.* **30B,** 455–464 (1996).

S. ÖZEKICI, “Average Waiting Times in Queues with Scheduled Batch Services,” *Transp. Sci.* **21,** 55–61 (1987).

S. ÖZEKICI AND S. ŞENGÖR, “On a Rail Transportation Model with Scheduled Services,” *Transp. Sci.* **28,** 246–255 (1994).

E. R. PETERSEN, “Over-the-Road Transit Time for a Single Track Railway,” *Transp. Sci.* **8,** 65–74 (1974).

E. R. PETERSEN, “Interference Delays on a Partially Double-Tracked Railway with Intermediate Signalling,” *J. Transp. Res. Forum* **16,** 55–62 (1975).

E. R. PETERSEN, “Railyard Modeling. Part I. Prediction of Put-Through Time,” *Transp. Sci.* **11,** 37–49 (1977a).

E. R. PETERSEN, “Railyard Modeling. Part II. The Effect of Yard Facilities on Congestion,” *Transp. Sci.* **11,** 50–59 (1977b).

E. R. PETERSEN AND A. J. TAYLOR, “A Structured Model for Rail Line Simulation and Optimization,” *Transp. Sci.* **16,** 192–205 (1982).

E. R. PETERSEN AND A. J. TAYLOR, “Line Block Prevention in Rail Line Dispatch and Simulation Models,” *INFOR* **21,** 46–51 (1983).

E. R. PETERSEN AND A. J. TAYLOR, “Design of a Single-Track Rail Line for High Speed Trains,” *Transp. Res.* **21A,** 47–57 (1987).

E. R. PETERSEN, A. J. TAYLOR, AND C. D. MARTLAND, “An Introduction to Computer Aided Train Dispatch,” *J. Adv. Transp.* **20,** 63–72 (1986).

C. E. PHILIP AND J. M. SUSSMAN, “Inventory Model of the Railroad Empty-Car Distribution Process,” *Transp. Res. Rec.* **656,** 52–60 (1977).

W. B. POWELL, “A Stochastic Model of the Dynamic Vehicle Allocation Problem,” *Transp. Sci.* **20,** 117–129 (1986).

W. B. POWELL, “An Operational Planning Model for the Dynamic Vehicle Allocation Problem with Uncertain Demands,” *Transp. Res.* **21B,** 217–232 (1987).

W. B. POWELL, “Making the Solution Fit the Problem,” *INFORMS Rail Applications Newsl.*, Spring 1995, 6–9.

W. B. POWELL, P. JAILLET, AND A. ODONI, “Stochastic and Dynamic Networks and Routing,” In *Network Routing,* M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser (eds.), North-Holland, Amsterdam, 141–295, 1995.

W. B. POWELL AND Y. SHEFFI, “Design and Implementation of an Interactive Optimization System for Network Design in the Motor Carrier Industry,” *Operations Res.* **37,** 12–29 (1989).

K. V. RAMANI, “An Information System for Allocating Coach Stock on Indian Railways,” *Interfaces* **11(3),** 44–51 (1981).

K. V. RAMANI AND B. K. MANDAL, “Operational Planning of Passenger Trains in Indian Railways,” *Interfaces* **22(5),** 39–51 (1992).

L. L. RATCLIFFE, B. VINOD, AND F. T. SPARROW, “Optimal Prepositioning of Empty Freight Cars,” *Simulation* **42,** 269–275 (1984).

R. E. RIVIER AND P. TZIEROPOULOS, “Interactive Graphics Models for Railway Operational Planning,” In *The Practice of Transportation Planning,* M. Florian (ed.), Elsevier Science Publishers, Amsterdam, 245–259, 1984.

R. E. RIVIER AND P. TZIEROPOULOS, “Computer-Aided Planning of Railway Networks, Lines and Stations,” In *Computers in Railway Management,* T. K. S. Murthy, L. S. Lawrence and R. E. Rivier (eds.), Computational Mechanics Publications, Berlin, 3–16, 1987.

F. J. M. SALZBORN, “Timetables for a Suburban Rail Transit System,” *Transp. Sci.* **3,** 297–316 (1969).

F. J. M. SALZBORN, “The Minimum Fleetsize for a Suburban Railway System,” *Transp. Sci.* **4,** 383–402 (1970).

R. L. SAUDER AND W. M. WESTERMAN, “Computer Aided Train Dispatching: Decision Support Through Optimization,” *Interfaces* **13(6),** 24–37 (1983).

H. D. SHERALI AND C. H. TUNCBILEK, “Static and Dynamic Time–Space Strategic Models and Algorithms for Multilevel Rail-Car Fleet Management,” *Management Sci.* **43,** 235–250 (1997).

M. W. SIDDIQEE, “Investigation of Sorting and Train Formation Schemes for a Railroad Hump Yard,” In *Traffic Flow and Transportation,* G. F. Newell (ed.), Elsevier, New York, 377–387, 1972.

M. E. SMITH, “Keeping Trains on Schedule: On-line Planning Systems for the Advanced Railroad Electronics System (ARES),” *J. Transp. Res. Forum* **31,** 17–24 (1990).

S. SMITH AND Y. SHEFFI, “Locomotive Scheduling under Uncertain Demand,” *Transp. Res. Rec.* **1251,** 45–53 (1988).

S. SPIECKERMANN AND S. VOSS, “A Case Study in Empty Railcar Distribution,” *Eur. J. Operational Res.* **87,** 586–598 (1995).

S. SUZUKI, “A Method of Planning Yard Pass Trains on a General Network,” In *Operational Research '72,* M. Ross (ed.), North Holland, Amsterdam, 353–361, 1973.

B. SZPIGEL, “Optimal Train Scheduling on a Single Track

Railway," In *Operational Research '72,* M. Ross (ed.), North-Holland, Amsterdam, 343–352, 1973.

M. A. THOMET, "A User-Oriented Freight Railroad Operating Policy," *IEEE Trans. Systems, Man Cybernet.* **1,** 349–356 (1971).

M. A. TURNQUIST AND M. S. DASKIN, "Queuing Models of Classification and Connection Delay in Railyards," *Transp. Sci.* **16,** 207–230 (1982).

C. D. VAN DYKE, "The Automated Blocking Model: A Practical Approach to Freight Railroad Blocking Plan Development," *Transp. Res. Forum* **27,** 116–121 (1986).

C. D. VAN DYKE, "Dynamic Management of Railroad Blocking Plans," *Transp. Res. Forum* **29,** 149–152 (1988).

W. W. WHITE AND A. M. BOMBERAULT, "A Network Algorithm for Empty Freight Car Allocation," *IBM Systems J.* **8,** 147–169 (1969).

M. B. WRIGHT, "Applying Stochastic Algorithms to a Locomotive Scheduling Problem," *J. Operational Res. Soc.* **40,** 187–192 (1989).

S. YAGAR, F. F. SACCOMANNO, AND Q. SHI, "An Efficient Sequencing Model for Humping in a Rail Yard," *Transp. Res.* **17A,** 251–262 (1983).

G. YU, (ed.), *Operations Research in the Airline Industry,* Kluwer Academic Publishers, Boston, MA, 1998.

K. ZIARATI, Affectation des Locomotives aux Trains, PhD dissertation, École Polytechnique de Montréal, Canada, 1997.

K. ZIARATI, F. SOUMIS, AND J. DESROSIERS, Locomotive Assignment using Train Delays, Technical report G-97-27, GERAD, École des Hautes Études Commerciales de Montréal, Canada, 1997.

K. ZIARATI, F. SOUMIS, J. DESROSIERS, S. GÉLINAS, AND A. SAINTONGE, "Locomotive Assignment with Heterogeneous Consists at CN North America," *Eur. J. Operational Res.* **97,** 281–292 (1997).

K. ZIARATI, F. SOUMIS, J. DESROSIERS, AND M. M. SOLOMON, A Branch-First, Cut-Second Approach for Locomotive Assignment, Technical report G-98-11, GERAD, École des Hautes Études Commerciales de Montréal, Canada, 1998.

P. J. ZWANEVELD, L. G. KROON, H. E. ROMEIJN, M. SALOMON, S. DAUZÈRE-PÉRÈS, S. P. M. VAN HOESEL AND H. W. AMBERGEN, "Routing Trains through Railway Stations: Model Formulation and Algorithms," *Transp. Sci.* **30,** 181–194 (1996).

(Received: February 1998; revisions received: July 1998; accepted: July 1998)