Original Article

Institution of MECHANICAL ENGINEERS

Journal of RAIL AND RAPID TRANSIT

Proc IMechE Part F:
J Rail and Rapid Transit
2019, Vol. 233(6) 629–639

DOI: 10.1177/0954409718804427
journals.sagepub.com/home/pif
SAGE

Transitioning from flexible to structured heavy haul operations to expand the capacity of single-track shared corridors in North America

C Tyler Dick, Darkhan Mussanov and Nao Nishio

Abstract

North American freight railroads typically exhibit flexible operations with variable train schedules where freight trains are dispatched when filled to length or weight capacity, and train dispatchers resolve train conflicts in real time, making it an ideal approach for low-cost transportation of bulk commodities. Recently, North American railways have experienced a substantial decline in demand for bulk transportation of coal while premium intermodal traffic that must operate on more structured schedules has reached record levels. On the predominantly single-track railway network in North America, operating a mixture of trains with differing schedule flexibility creates fluctuating patterns of train meets that increase train delay, constrain capacity, and decrease utilization of capital-intensive track infrastructure. This study investigates the relationship between line capacity, amount of schedule flexibility, level of service, and the proportion of flexible and scheduled trains operating on a rail line. Using the Rail Traffic Controller software, different combinations of these factors are simulated to determine the capacity of representative North American rail corridors with heavy haul and priority intermodal traffic. The simulation results suggest that introducing more structured schedules for heavy haul traffic is an effective alternative to investment in additional second main track infrastructure to gain sufficient capacity for growing scheduled intermodal traffic on traditional bulk commodity freight corridors.

Keywords

Railway operations, simulation, freight transportation, single track, mainline capacity, train delay

Date received: 10 May 2018; accepted: 20 August 2018

Introduction

Heavy haul railways provide the safe, efficient, and reliable transportation of freight with particular applications to bulk commodities. Heavy haul railway operations often use dedicated unit or shuttle trains to gain efficiencies by maximizing the quantity of freight transported by each train and continuously cycling locomotives and railcars between loading and unloading terminals.[1] To achieve these efficiencies, the initial terminal departure or final arrival times of heavy haul freight trains are often dictated by bulk freight commodity production plans and vessel sailing schedules instead of a predetermined railway timetable. Bulk commodity trains depart when an entire trainload of freight is ready to be shipped and a train crew and equipment are available.[2] Many heavy haul operations exhibit schedule flexibility as trains may depart terminals at random times through the day or week. Without a fixed timetable and preplanned locations for each train meet on single track, rail traffic controllers must resolve train conflicts in real time under flexible operations. This paper will demonstrate that flexible heavy haul operations consume more capacity than structured operations where trains depart within shorter time windows, and introducing more structured schedules for heavy haul traffic can effectively replace investment in additional second main track as a means to increase railway line capacity.

Many North American rail lines once dominated by heavy haul bulk commodity traffic operating on flexible schedules are now seeing increased volume of intermodal trains transporting shipping containers and highway trailers that require more structured

Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, USA

Corresponding author:
C Tyler Dick, 1241 Newmark Civil Engineering Lab, MC-250 205 N. Mathews Avenue, Urbana, IL 61801, USA.
Email: ctdick@illinois.edu

operations to meet shipper needs. Following several years of traffic declines due to economic recession, freight rail traffic in the United States has slowly returned to the peak levels set in 2006. However, the composition and geographic distribution of US rail traffic have substantially changed since 2006. Coal traffic has declined by over 20% while intermodal traffic has set new records, with particularly strong growth in domestic intermodal traffic requiring predictable service on precise schedules to be competitive with the highway mode.[3]

Heavy haul bulk commodity and intermodal train types have unique level-of-service (LOS) requirements and different train speed, priority, length, and weight characteristics. Allowing for schedule flexibility to maintain the efficiency of bulk commodity trains while simultaneously providing the reliable LOS required by intermodal trains presents a substantial operational challenge. Although constructing additional segments of second main track may provide sufficient capacity for these two distinct operations to coexist on the same single-track mainline, it may also be possible to maximize the capacity and LOS of these rail lines by finding a balance between flexible and structured operations for heavy haul traffic.

This study investigates the relationship between mainline capacity, train delay (LOS), amount of schedule flexibility, and the combination of flexible and scheduled trains operating on a rail line. Simulation experiments are used to calculate the capacity of a representative North American single-track mainline under operating conditions ranging from complete flexibility with random train departures, to highly structured operations according to a precisely planned timetable intended to minimize train delay. The conclusions of this research provide a better understanding of the interaction between line capacity and schedule flexibility, and the equivalency between schedule flexibility and expanded mainline track infrastructure. Railway practitioners can use this knowledge to make better-informed decisions regarding mainline capacity and the amount of flexibility to include in heavy haul train operating plans.

Background

Flexible and structured railway operations

Most North American freight trains do not operate according to a prescribed timetable; although the general pattern of train operations on a given day of the week may be similar from week to week, a specific train may depart a terminal at different times each day. In cycling between bulk commodity loading and unloading terminals, heavy haul freight trains may randomly enter the rail network according to bulk commodity production and shipping schedules. This style of train operation is referred to as "improvised" or "flexible operation."[4] Flexible operations allow terminals to adapt to changing traffic conditions and shipment demand by changing train departure times.[5] Terminals may hold trains past their departure times to maximize the freight hauled by each train and preserve the efficiency of the heavy haul operation.[6]

The flexible style of railway operations is in contrast to "structured operations" where a detailed timetable specifies preplanned departure times, exact locations to resolve meet and pass conflicts with other trains, and arrival times at intermediate points on the route and the final destination terminal.[4] The structured style of railway operations is common in Europe and Asia, and also on commuter rail and transit systems in North America. In Europe, the structured method of operation applies to all train types, including heavy haul freight trains.[7]

This study uses the term "schedule flexibility" to describe the range of actual terminal departure times relative to a baseline train operating plan with target departure times. Low schedule flexibility corresponds to more structured operations where all trains depart relatively close to their planned departure times. A fully structured operation exhibits no schedule flexibility; all trains depart exactly at their planned time. When train departures are distributed over a wider range about the planned departure time, the operation has high schedule flexibility. A fully flexible operation features trains that depart terminals randomly over each 24 h period of operations.

Prior research work

North American heavy haul railway operators define line capacity as the largest volume of traffic (in trains per day) that can be operated over a mainline segment while meeting a required LOS.[8] Railway performance metrics are used to establish the required LOS, with maximum allowable average train delay being the most common. In this context, train delay is defined as the difference between the actual train running time over a mainline segment and the minimum running time of the same train without interference from conflicts with other trains. The actual running time is the total time elapsed while a train traverses the segment in the presence of train traffic, including the time the train is in motion and the time the train is stopped waiting for other trains or unexpected sources of delay. By this definition, train delay includes the time a train is stopped waiting for other trains. This definition is different from other international definitions of train delay where scheduled time spent dwelling in stations or for planned trains meets is not considered as delay.[9] The average train delay for all trains on a given mainline can be used to calculate line capacity, or the specific LOS requirements for a certain train type may also be used to establish line capacity.[10]

Delay-based definitions of line capacity have been formalized into parametric capacity models, such as those developed by the US Federal Railroad Administration[11] and Canadian National Railway,[8] that describe relationships between line capacity, train delay, and various track infrastructure layout, traffic, and operating parameters. Neither model explicitly considers the amount of schedule flexibility in determining line capacity.

Martland[12] observed that heavy haul line capacity and terminal operating plans need to consider disruptions and operational variability. Martland also suggested that changes to operations may allow railroads to increase capacity without investments in mainline track infrastructure expansion. Ede et al.[13] investigated the balance between parameters involving operations, train characteristics, route infrastructure investment, and maintenance scheduling in establishing the most cost-effective heavy haul operation. Their research suggested that bunching train departures due to variability reduces operational capacity compared to maintaining evenly spaced departures, and that sharing heavy haul routes with other train types of trains created capacity challenges for the heavy haul operation. This paper attempts to quantify the qualitative observations made by these researchers on the effects of departure variability and train-type interactions on line capacity.

Dingler et al.[14] investigated the capacity effect of interactions between lower speed heavy haul bulk unit trains and priority higher speed intermodal trains on a representative single-track mainline. The largest train delays were observed when the volume of unit and intermodal trains on the simulated mainline were equal. Dingler did not consider the unique schedule flexibility characteristics of each train type. Subsequent heterogeneity research included randomized train departures but did not include schedule flexibility as a factor in the experimental design.[15,16]

Boysen[17] developed a model to examine how operating parameters, train characteristics, and stakeholder needs can influence railway line capacity. The model suggested that decreasing train heterogeneity could increase capacity but schedule flexibility is not explicitly addressed. A subsequent study investigated the heavy haul capacity consumed by allocating timetable slots to passenger trains and maintenance activities on a heavy haul route transporting iron ore in northern Norway and Sweden.[18] For passenger trains to adhere to the timetable, iron ore train operations are constrained to preplanned timetable slots. To provide the flexibility in heavy haul train departures needed to meet fluctuations in iron ore production and shipment demand, the freight operator purchases additional timetable slots above the minimum number required to transport the average daily volume of iron ore.[19] Even if they are otherwise ready to depart, trains are held at end terminals until their scheduled departure time and many timetable slots are not used during periods with low demand.

The line capacity of two different control systems under structured and flexible operations was compared through a Dutch case study.[20] To determine the capacity of structured operations, the standard UIC compression approach was used while Monte Carlo simulation and rescheduling algorithms were required for flexible operations. The purpose of the study was to investigate the capacity of each control system to accommodate unplanned events and unscheduled trains as opposed to the specific capacity effects of varying amounts of schedule flexibility.

This paper builds upon a previous study of the relationship between train delay and different levels of schedule flexibility under homogeneous operations and a constant traffic volume on representative mainlines in North America.[21] Beginning with an initial timetable designed to minimize average train delay, introducing schedule flexibility by increasing the range of departure time variability caused average train delay to increase. The relationship between schedule flexibility and train delay was found to be nonlinear: low levels of schedule flexibility resulted in disproportionately large increases in train delay while, beyond a certain level of schedule flexibility, additional incremental train delay was not observed as schedule flexibility reached high levels. Additional investigation revealed that introducing schedule flexibility increases runtime and decreases train velocity relative to structured operations.[22] Subsequent research examined the relative train delay experienced by various combinations of flexible and scheduled trains traversing the same mainline at different levels of schedule flexibility.[23] The experimental results indicate that, particularly as line capacity utilization increases, scheduled trains have difficulty maintaining a high LOS when operating among flexible trains.

Hypothetical relationship

This study investigates schedule flexibility and its relationship to the capacity of representative North American single-track mainlines and the LOS (train delay) experienced by particular train types for certain traffic mixtures composed of flexible and scheduled freight trains. This study also aims to demonstrate that investments in additional mainline track infrastructure can be avoided and equivalent capacity gained by reducing the level of schedule flexibility of heavy haul freight train operations.

Previous research by Dick and Mussanov[21] indicated that when operating with schedule flexibility at a constant traffic volume, the incremental increase in train delay diminished as schedule flexibility reached higher levels (Figure 1(a)). On North American single-track mainlines, it is generally accepted that increasing traffic volume leads to exponential increases in freight train delay (Figure 1(b)).[8,14] The hypothesized set of delay–volume curves in Figure 1(b) can be deduced from the form of the relationship between

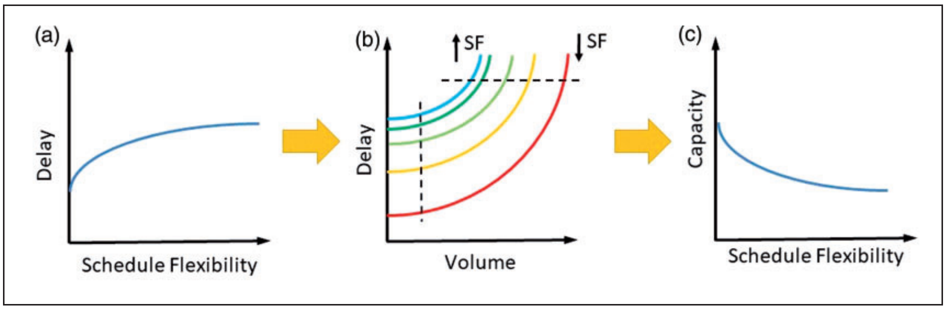

Figure 1. Relationships between (a) schedule flexibility and train delay for constant volume and (b) volume and train delay as a function of schedule flexibility lead to (c) hypothetical relationship between schedule flexibility and line capacity for given maximum train delay. SF: schedule flexibility.

schedule flexibility and delay for a constant volume in Figure 1(a). The delay–volume curves (Figure 1(b)) are more closely spaced when schedule flexibility is higher to reflect the decreasing sensitivity of delay to additional schedule flexibility (Figure 1(a)). This paper describes the results of simulation experiments designed to determine if this hypothetical relationship is present on representative single-track mainlines.

The form of Figure 1(b) suggests that the same average train delay may be exhibited by an operation with higher traffic volume and lower schedule flexibility, or an operation with a lower traffic volume and greater schedule flexibility. By setting a maximum allowable train delay (LOS) to define line capacity, the corresponding combinations of volume and schedule flexibility in Figure 1(b) produce a hypothetical relationship between line capacity and schedule flexibility for a given LOS (Figure 1(c)). As schedule flexibility decreases from high levels toward more structured operations, line capacity is expected to increase. The additional line capacity obtained under fully structured operations (zero schedule flexibility) corresponds to the capacity lost by allowing some (or all) trains to operate on flexible schedules. Based on previous research on the incremental capacity of transitioning from single to double track,[16] the capacity gained by decreasing schedule flexibility can be equated to a savings in second main track infrastructure investment that would have been required to provide the same capacity under flexible operations.

Methodology

Rail Traffic Controller (RTC)

This study uses RTC, the industry-leading rail traffic simulation software in the United States, to calculate train delay and capacity metrics via simulation. Specially developed for the flexible North American railway operating environment, RTC is used by most US Class I railroads, Amtrak, and various rail operations consultants.[24] RTC simulates the movement of trains over mainlines with specific route characteristics and emulates the decisions made by rail traffic controllers in resolving conflicts between trains. Users can alter different infrastructure, train, and control system parameters within the RTC simulation to observe and analyze the corresponding train delay output.

Under structured operations, scheduled trains depart their originating terminals according to a timetable specified in the RTC input. Flexible trains randomly depart their originating terminal earlier or later than the departure time specified in the train plan according to a user-specified range. If a schedule flexibility of 60 min is assigned to a particular train, RTC will randomly depart that train as early as 60 min before or 60 min after departure time in the train plan. Within the 120 min range of possible departure times, the exact departure time for each simulated day will vary according to a uniform distribution. In the absence of knowledge regarding the actual distribution of train departure times, this research uses the uniform distribution because it is already built in to the randomization functions within the RTC software. Users may specify a different amount of schedule flexibility for each train in the RTC simulation, allowing RTC to simulate a representative North American freight mainline under different combinations of heavy haul bulk commodity trains with schedule flexibility and high-priority intermodal trains with scheduled departure times.[25]

To calculate average train delay for each type of train in a simulation scenario, the train delay output from RTC for a given train type is averaged across all simulation days and normalized by the respective total train-mile (train-km) accumulated by trains of that type.

Under flexible operations, the exact train departure pattern is different for each simulated day. To capture the effect of different departure patterns and calculate a robust average train delay response, simulation scenarios are replicated. The required number of replicates was estimated by replicating one scenario in the experiment design 100 times with different initial random seeds. The average train delay response was observed to stabilize after seven replications. On this basis, each experiment design scenario is simulated for five days of train operations and then replicated 10 times. For each

Table 1. Baseline schedule, mainline route, and train parameters.

Parameter	Value
Route length	240 mile (386 km)
Siding spacing	10 mile (16 km)
Siding length	2 mile (3.22 km)
Number of sidings	23
Scheduled departure interval	2 h
Maximum speed	23 mph (37 km/h)
Traffic volume	36 trains per day
Operating protocol	2-block, 3-aspect Centralized Traffic Control
Train consist	115 railcars at 125 t each 1.29 mile (2.07 km) total length
Locomotive type	SD70 4300 hp (3206 kW) 3 locomotives per train

experimental scenario, this approach produces 50 days of mainline train delay information that is processed into a single average train delay value representing those experimental conditions.

Baseline scheduled train plan

Consistent with previous research work on this topic,[25] a combination of baseline schedule and track infrastructure layout was developed to minimize train delay for 36 trains per day on a single-track mainline (Table 1). The "return-grid" operating model was used to design the baseline schedule such that when trains alternately depart each end terminal at even intervals, every train meet conflict occurs at one of the evenly spaced passing sidings (passing loops). Since all of the train conflicts are planned to occur at passing sidings, the overall delay created by these train meets is minimized. The baseline schedule is subsequently perturbed to introduce schedule flexibility and additional traffic volume to produce the range of scenarios in the experimental design. The mainline segment is 240 mile (386 km) long with 23 passing sidings (passing loops) that are each 2 mile (3.22 km) long and placed every 10 mile (16 km) on-center. Every train has an identical consist of 125 railcars that are representative of North American heavy haul freight operations.

Experimental design

The different simulation scenarios in the main experimental design include constant route infrastructure and three variable factors: traffic volume, schedule flexibility, and traffic composition (Table 2). The combination of factors and factor levels in the main experimental design produces 150 different scenarios for simulation with RTC. To extend the main experimental design, an additional set of simulations was conducted with 42 trains per day, 720 min of schedule flexibility, and an increasing amount of added second mainline track infrastructure (in the form of double-track segments connecting existing passing sidings). These supplementary simulations quantify the amount of track infrastructure investment required under flexible operations to obtain the equivalent capacity created by reducing the number of flexible trains and their associated schedule flexibility.

Table 2. Experimental design factors and factor levels.

Factor	Factor levels
Traffic volume (trains per day)	24, 28, 32, 36, 40, 44
Schedule flexibility ($\pm$min)	0, 10, 60, 120, 720
Traffic composition (percent of trains with flexible schedules)	0, 25, 50, 75, 100
Amount of double track (%)	18, 25, 32, 38, 45, 59

To change the traffic volume from the initial baseline schedule of 36 trains per day, trains were removed and added in pairs to maintain directional balance and the structure of the "return-grid" schedule. For example, when four trains are removed from the initial 36-train schedule to achieve a volume of 32 trains per day, to preserve the ideal "return-grid" schedule, the remaining 32 trains retain their original departure time slots and are not respaced to depart at even intervals. To increase the traffic volume above 36 trains per day, an appropriate number of departure time slots used by one train in the baseline schedule are converted to depart two successive trains (departing at the minimum headway allowed by the block signal system) in the train plan at the higher traffic volume.

Two types of trains comprise the simulated mainline operation: scheduled and flexible. Scheduled trains adhere to the specific terminal departure times in the baseline schedule irrespective of the schedule flexibility factor level. The initial terminal departure times of flexible trains are randomly selected from a uniform distribution over a time range defined by the schedule flexibility factor and centered on the planned departure time specified in the baseline schedule. For example, 60 min of schedule flexibility defines a uniform distribution of departure times over a 120 min range extending 60 min before the planned departure time and up to 60 min after. Five levels of the schedule flexibility factor are considered: 0, 10, 60, 120, and 720 min. Zero minutes of schedule flexibility represent the structured operation on the baseline schedule, and 720 min defines fully random departures over each 24 h period. With previous research concluding that moving from medium to high levels of schedule flexibility does not substantially change train delay,[21] additional factor levels between 120 and 720 min of schedule flexibility were not included in the experimental design.

To investigate if the relationship between line capacity and schedule flexibility is sensitive to the mixture of flexible and scheduled trains, the experimental design includes different traffic compositions. Factor levels of 0, 25, 50, 75, and 100% flexible trains define the fraction of trains that operate with flexible schedules at a particular traffic volume. For example, at 32 trains per day and 25% flexible trains, the simulation consists of eight flexible trains per day operating at the specified level of schedule flexibility and 24 scheduled trains per day adhering to precise departure time. Higher priority is assigned to the scheduled trains within RTC.

Results

Schedule flexibility and line capacity

For a given traffic composition and level of schedule flexibility, the normalized average train delay output from the simulation scenarios exhibits the expected relationship with traffic volume (Figure 2(a)). Consistent with previous research,[21] at a constant traffic volume, increasing schedule flexibility leads to increases in train delay.

Since North American railways typically estimate line capacity on the basis of train delay, the delay–volume curves in Figure 2(a) can be transformed to create a relationship between schedule flexibility and line capacity for a given traffic composition and required LOS (Figure 2(b)). Figure 2(a) and (b) illustrates the specific relationship between line capacity and schedule flexibility obtained when the minimum LOS for all trains is set to a maximum permissible average train delay of 40 min per 100 train-mile (160 train-km).

The general trends in Figure 2(a) and (b) are consistent with the relationships hypothesized in Figure 1(b) and (c). At the 40 min LOS and 75% flexible trains traffic composition illustrated in Figure 2(b), transitioning the flexible trains from fully flexible to structured operations increases line capacity by approximately four trains per day.

Regression model for performance of individual train types

The capacity estimates in the previous section are made on the basis of average train delay of all trains and do not consider LOS requirements specific to a certain train type. Using techniques documented by Shih et al.,[10] train-type performance was included in a more comprehensive regression model describing the capacity of the simulated mainline (equation (1)). The regression model includes schedule flexibility, traffic volume, and traffic composition (percent

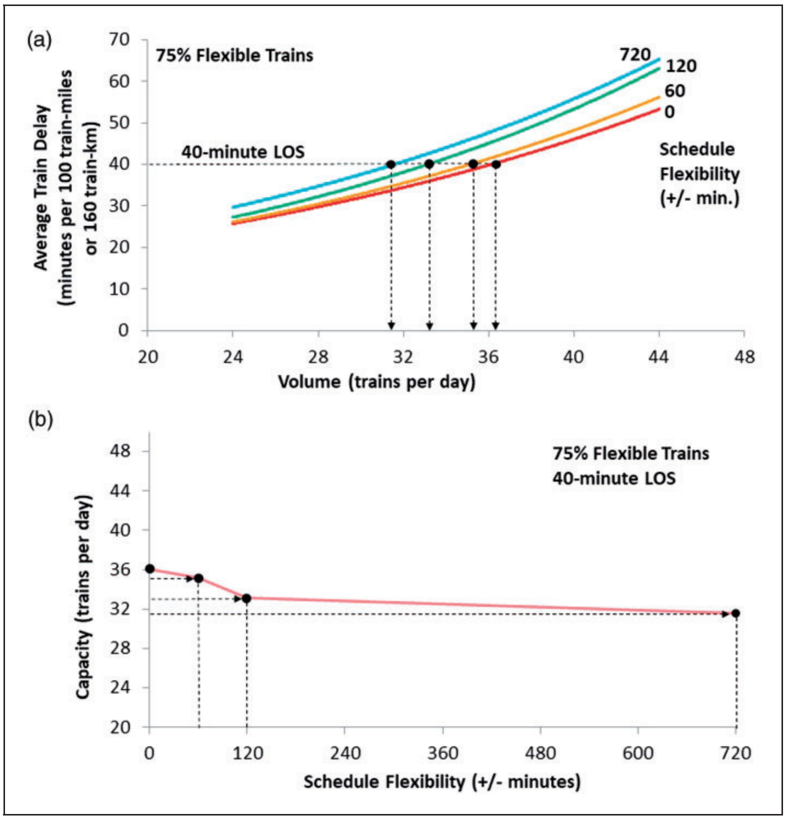

Figure 2. (a) Average train delay and traffic volume for different levels of schedule flexibility transformed into (b) line capacity for a given schedule flexibility and maximum allowable train delay. LOS: level of service.

flexible trains) as independent variables, and train delay specific to scheduled and flexible trains as independent variables. The model has an R-squared of 0.94 and significant interactions with p-values below 0.01, making it a good predictor of the train-type-specific train delay across the range of simulated factor levels

$$D_t = f_t(c, SF) * V^2 + g_t(c, SF) * V + h_t(c, SF) \quad (1)$$

where:

D_t = average normalized train delay for train type t
c = traffic composition (percent flexible trains)
SF = schedule flexibility
V = traffic volume (trains per day)
g_t, f_t, h_t = functions representing the first- and second-order delay–volume relationship specific to train type t

To transform this delay model into an equation for line capacity, equation (1) is set to equal the maximum allowable average train delay for each train type. The resulting expression can be solved for traffic volume using the quadratic formula (equation (2)). The traffic volume V* calculated with this equation is the capacity of the simulated mainline as defined by the LOS required by a particular type of train

$$V^* = \frac{-g_t(c, SF) + \sqrt{g_t^2(c, SF) - 4 * f_t(c, SF) * (h_t(c, SF) - D_t)}}{2 * f_t(c, SF)} \quad (2)$$

where:

D_t* = maximum allowable average normalized train delay for train type t
V* = line capacity to ensure train type t meets required LOS

Using equation (2), the capacity can be estimated for a range of schedule flexibility and traffic composition values (Figure 3). The capacity values in the figure were calculated using 21 min of delay per 100 train-mile (160 train-km) for scheduled trains and 38 min of delay for flexible trains as the specific LOS requirements for each train type. Since the transformation process is performed once for each train type using its specific maximum allowable train delay value, each traffic composition has two different capacity curves, with each capacity curve defined by one train-type-specific LOS. The final capacity at a particular level of schedule flexibility is the lesser of the two capacity values for that combination of schedule flexibility and traffic composition.

The relative values of capacity determined by the scheduled and flexible train LOS at a given schedule flexibility (Figure 3) suggest the LOS requirement for scheduled trains defines overall line capacity at lower values of schedule flexibility. As schedule flexibility increases, overall line capacity is determined by the flexible train LOS requirement. As indicated by the slope of the capacity contours and range of capacity values defined by each train type, scheduled trains are less sensitive to changes in schedule flexibility compared to flexible trains.

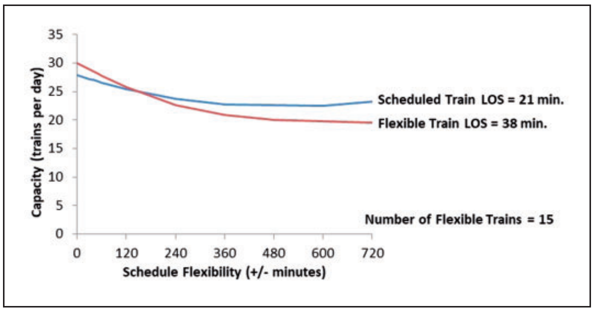

Figure 3. Capacity contours defined by scheduled and flexible train LOS. LOS: level of service.

From the practitioner perspective, these results suggest that line capacity is constrained by the flexible train LOS if externalities and disruptions necessitate flexible operations. If the route can shift toward structured operations, line capacity increases and is limited by the scheduled train LOS.

Capacity of different traffic compositions

Comparing traffic mixtures reveals the relative capacity lost or gained from operating different combinations of scheduled and flexible trains across a range of schedule flexibility (Figure 4(a) and (b)).

Comparing traffic compositions, regardless of the train type LOS used to define line capacity, the highest capacity is achieved when there are fewer flexible trains and less schedule flexibility. The lowest line capacity is observed when equal numbers of scheduled and flexible trains comprise the traffic on the route, consistent with the observations of previous research on rail traffic heterogeneity. Scenarios with 75% flexible trains have slightly higher capacity than those with 50% flexible trains. This somewhat counterintuitive result arises because, when a majority of the trains are flexible and assigned low priority, it is easier for the small number of high-priority scheduled trains to maintain their required LOS; the elevated status of the remaining scheduled trains compensates for the additional randomness created by the large number of meets between flexible trains. When the traffic composition shifts such that a majority of the trains are scheduled (i.e. $<$ 50% flexible trains), the large number of scheduled trains diminishes the priority effect but there is less randomness, resulting in many low-delay preplanned train meets between scheduled trains that increase capacity overall. For the presented scenarios, changing the traffic composition at a

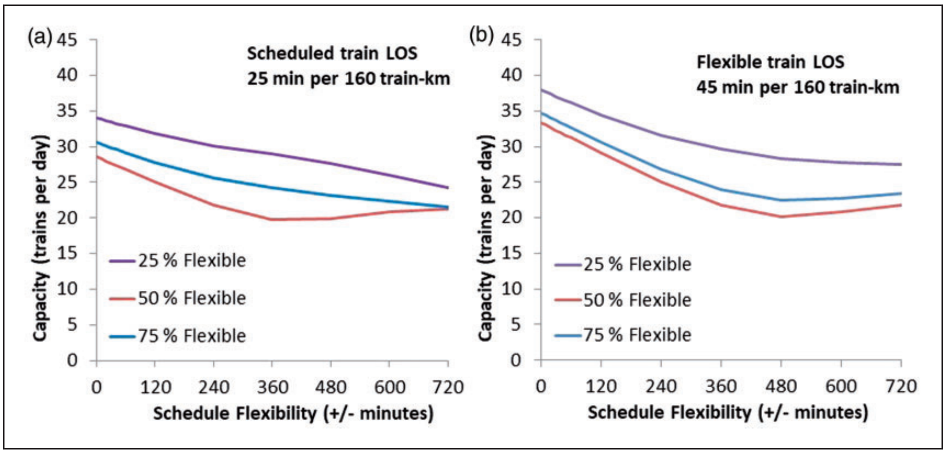

Figure 4. Capacity under different traffic compositions and schedule flexibility as defined by (a) flexible train LOS and (b) scheduled train LOS. LOS: level of service.

constant level of schedule flexibility can alter the line capacity by as much as approximately five trains per day.

Train-type LOS requirements and line capacity

To demonstrate the relative LOS performance of different train types, the capacity for a traffic composition of 50% flexible trains was plotted across a range of schedule flexibility for various train-type-specific LOS (Figure 5). Certain combinations of flexible and scheduled train LOS yield the same line capacity. As illustrated in the figure, a 65 min flexible train LOS yields a capacity relationship that, with $p = 0.05$, belongs to the same population as the capacity curve defined by a 45 min scheduled train LOS. On average, the flexible trains will have 20 min more delay per 100 train-mile (160 train-km) compared to the scheduled trains. To obtain at most a 45 min LOS for both train types, line capacity would decrease substantially to that defined by the 45 min flexible LOS contour and the scheduled trains would only experience 25 min of train delay on average.

This difference in line capacity created by train-type LOS requirements reinforces the need to consider the performance of individual types of trains when calculating mainline capacity. If premium scheduled intermodal trains are more sensitive to delays than flexible heavy haul trains, the specific LOS experienced by the scheduled trains will be a better metric for establishing line capacity than the overall average train delay. Evaluating railway line capacity by the average train delay over all types of trains does not guarantee the performance of any particular train type; low delay for certain train types may only come at the expense of high delay to other train types.

Volume of flexible trains

In practice, a heavy haul operator can reduce both the schedule flexibility and the number of flexible trains when transitioning from flexible to structured operations. The combinations of schedule flexibility and flexible train volumes corresponding to a given average train delay (LOS) define a relationship between the number of flexible trains and line capacity (defined by a given LOS) for various levels of schedule flexibility (Figure 6).

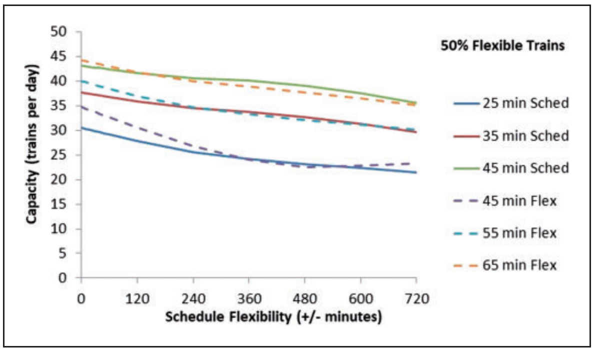

Figure 5. Capacity for various train-type-specific LOS and 50% flexible trains.

Under structured operations, the simulated single-track mainline has a capacity of 41 scheduled trains per day when the average LOS of all trans is fixed at 40 min of train delay per 100 train-mile (160 train-km). To replace eight of the scheduled trains with flexible trains operating at 120 min of schedule flexibility, the capacity of the line must be reduced to approximately 35 trains per day to maintain the required LOS. Alternatively, if the schedule flexibility of those eight trains can be limited to 60 min, the capacity only decreases to 37 trains per day. A capacity of 35 trains per day can also be obtained with 720 min of schedule flexibility but only if at most four of these highly flexible trains are operated on the line. It is possible to increase line capacity by moving to structured operations for most trains but still operating a small number of trains with high schedule flexibility. Adjusting both the number of flexible trains and their schedule flexibility provides practitioners with more

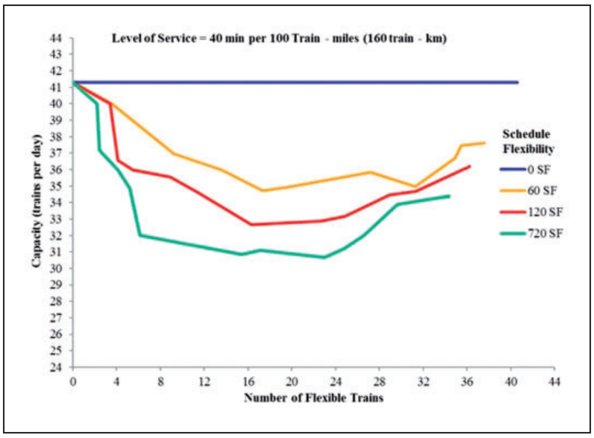

Figure 6. Number of flexible trains and line capacity at various levels of schedule flexibility for 40 min average LOS of all trains. SF: schedule flexibility.

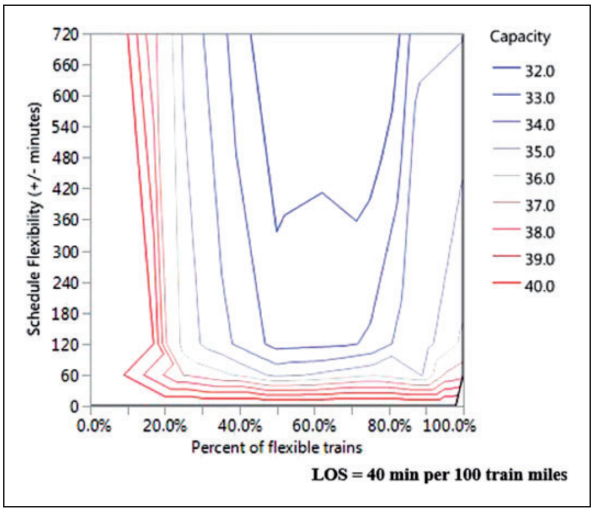

Figure 7. Percent of flexible trains and schedule flexibility providing a given line capacity. LOS: level of service.

options to maximize line capacity via structured operations while still accommodating the flexible schedules required by certain heavy haul trains.

The increasing magnitude of capacity that can be gained by transitioning from higher levels of schedule flexibility is further illustrated in Figure 6. Achieving a specific increase in capacity requires more flexible trains to change into scheduled trains when operating at higher schedule flexibility compared to operations with lower schedule flexibility.

Replotting the data to show contours of constant capacity within the space defined by traffic mixture (percent flexible trains) and schedule flexibility (Figure 7) further demonstrates the equivalent capacity of a small number of highly flexible trains and a larger number of less flexible trains.

The shape of the capacity contours in Figure 7 suggests the operational change that will most efficiently increase capacity depends on the current traffic mixture and amount of schedule flexibility. At low schedule flexibility, capacity is most efficiently increased by further decreases in the amount of schedule flexibility while changing the traffic mixture has little effect on capacity. When the traffic mixture includes less than 40% flexible trains and the schedule flexibility exceeds $\pm$120 min, capacity is most efficiently increased by further reductions in the number of flexible trains and less by reductions in schedule flexibility. Combined with practical constraints on how much schedule flexibility can be reduced and how many flexible trains can become scheduled, this finding suggests there is no single best pathway to increase capacity by transitioning from flexible to structured heavy haul freight operations.

Equivalent capacity of structured operations and infrastructure expansion

Although some decrease in schedule flexibility may be achieved through improved operating discipline, as described by Dick and Mussanov,[21] reducing schedule flexibility and transitioning traffic to a structured operation to increase capacity may not be without cost. To reduce schedule flexibility caused by unavailability of resources, railway operators may need to invest in additional crew, locomotives, and rolling stock. Similarly, to reduce schedule flexibility from failure-related disruptions, track, rolling stock, and control systems may require additional maintenance. If the move to structured operations is mainly driven by capacity and not to just improve service, investments required to reduce schedule flexibility can be justified by the potential increase in revenue from additional traffic and the relative cost of other strategies to increase capacity such as track infrastructure expansion. The capacity gained through operational changes to decrease schedule flexibility may allow railways to defer large investments in additional second main track that would otherwise be required to increase capacity.

To illustrate the magnitude of track infrastructure investment that can be saved by gaining the equivalent capacity via a transition to structured operations, one traffic composition was simulated on the baseline route with increasing amounts of second main track. As shown in Figure 6, for a LOS of 40 min, the simulated single-track mainline has a capacity of 30 trains per day (6 scheduled and 24 flexible, or 80% flexible trains) at $\pm$ 720 min of schedule flexibility. If the route transitions to a fully structured operation (zero percent flexible trains and zero schedule flexibility), the baseline single-track mainline has a capacity of 42 trains per day. By making the necessary operational changes to reduce schedule flexibility, capacity can be increased by 12 trains per day without investing in additional mainline infrastructure.

If the operator attempted to operate 42 trains per day on the baseline single track at the original traffic composition of 80% flexible trains (8 scheduled and

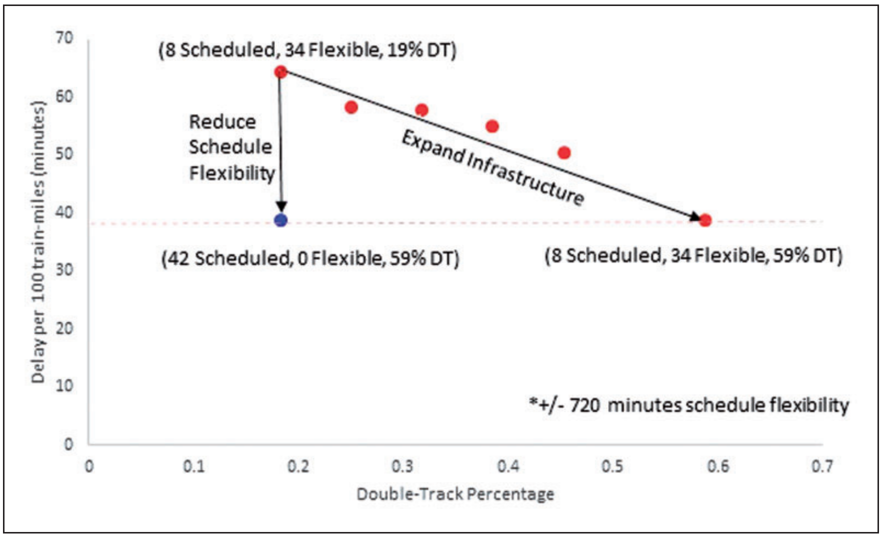

Figure 8. Capacity and LOS equivalency of infrastructure investment and transitioning to scheduled operations. DT: double track.

34 flexible) and ±720 min of schedule flexibility, the normalized train delay would be over 60 min per 100 train-mile (161 train-km) (Figure 8). Since this average train delay exceeds the required LOS, the line would be overcapacity.

To reduce the train delay under these operating conditions back to the original 40 min LOS and obtain equivalent capacity to structured operations, the mainline infrastructure must be expanded from 19 to 59% second main track by adding sections of double track between the existing passing sidings (Figure 8). Along the simulated 240 mile (386 km) mainline, 40% second main track is equivalent to 96 mile (155 km) of new track construction that can be avoided or deferred by transitioning to structured operations. If the simulated mainline was in an urban area or difficult terrain where numerous bridges and tunnels may be required, several hundred million US dollars of infrastructure investment could be saved by increasing capacity through a transition from flexible to structured operations. Practitioners can engage in a similar approach to quantify the benefits and costs of achieving a desired amount of line capacity through a combination of reducing schedule flexibility and expanding mainline track infrastructure.

Conclusion

This study uses RTC simulation to analyze the relationship between line capacity, schedule flexibility, traffic volume, and traffic composition for representative North American single-track railway corridors. For a given LOS and a constant track infrastructure layout, different traffic compositions of scheduled and flexible trains exhibit the trend of increasing capacity with decreasing schedule flexibility. The largest capacity gains are made when moving from low levels of schedule flexibility to completely structured operation. Mainlines operating at high levels of schedule flexibility do not experience increased capacity until schedule flexibility is substantially decreased to near-structured operations. The number of flexible trains and level of schedule flexibility can both be altered in combination to increase line capacity while still accommodating the flexible schedule needs of selected heavy haul trains. Although scheduled trains are still subject to train delay when a small number of highly flexible trains remain on the corridor, the average delay experienced by the scheduled trains will typically be lower than the average flexible train delay. In defining line capacity, it is important to consider the specific LOS required by each type of train and not just the average delay over all scheduled and flexible trains.

The capacity gained in transitioning from fully flexible to structured operations on a single-track mainline can also be obtained through equivalent investment in capital projects to expand mainline track infrastructure. Depending on the exact route and traffic conditions, the amount of new passing siding and second main track construction that can be avoided or deferred can be substantial and justify cost increases required to ensure that trains adhere to precise schedules. In particular, transitioning from flexible to structured heavy haul operations may prove to be a viable means to increase capacity in bottleneck situations where the remaining single-track segments on a mainline are infeasible to double track due to costly or insurmountable environmental or engineering obstacles.

Acknowledgements

The authors thank Eric Wilson and Berkeley Simulation Software, LLC for the use of Rail Traffic Controller simulation software. The authors also thank Mei-Cheng Shih, Graduate Research Assistant at the University of Illinois at Urbana-Champaign, for technical insight.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the National University Rail Center (NURail), a US DOT OST Tier 1 University Transportation Center, and the Association of American Railroads.

ORCID iD

C Tyler Dick http://orcid.org/0000-0002-2527-1320

References

1. Starr JT. *Evolution of the unit train: 1960–1969*. Chicago, IL: University of Chicago, 1976.
2. Grimes GA. Planning for unit train operations on Missouri Pacific. *J Transp Res Forum* 1981; 22: 121–126.
3. Association of American Railroads. *Freight railroad capacity and investment*. Washington, DC: Association of American Railroads, 2016.
4. Martland CD. Improving on-time performance for long-distance passenger trains operating on freight routes. *J Transp Res Forum* 2010; 47: 63–80.
5. Kwon O. *Service differentiation in railroad freight service*. PhD Thesis, Massachusetts Institute of Technology, USA, 1994.
6. Dong Y. *Modeling rail freight operations under different operating strategies*. PhD Thesis, Massachusetts Institute of Technology, USA, 1997.
7. Furtado F. U.S. and European freight railways: the differences that matter. *J Transp Res Forum* 2013; 52: 65–84.
8. Krueger H. Parametric modeling in rail capacity planning. In: *The winter simulation conference*, Phoenix, Arizona, 1999, pp.1194–1200. New York: Institute of Electrical and Electronics Engineers (IEEE).
9. Pouryousef H, Lautala P and White T. Railroad capacity tools and methodologies in the U.S. and Europe. *J Modern Transp* 2015; 23: 30–42.
10. Shih MC, Dick CT and Barkan CPL. Impact of passenger train capacity and level of service on shared rail corridors with multiple types of freight trains. *Transp Res Rec J Transp Res Board* 2015; 2475: 63–71.
11. Prokopy JC and Rubin RB. *Parametric analysis of railway line capacity*. Report, FRA, USA, 1975.
12. Martland CD. Modeling the effects of heavy haul freight service on terminal productivity. In: *Proceedings of International Heavy Haul Association, implementation of heavy haul technology for network efficiency, a special technical session*, Dallas, United States, 5–9 May 2003.
13. Ede BM, Polivka A and Tunna J. Optimizing capacity by trade-offs among train control, infrastructure, mechanical, and engineering considerations. In: *Proceedings of the 2007 IHHA specialist technical session*, Kiruna, Sweden, 2007, pp.413–421. International Heavy Haul Association (IHHA).
14. Dingler MH, Koenig A, Sogin SL, et al. Determining the causes of train delay. In: *Proceedings of the annual AREMA conference*, Orlando, United States, 29 August–1 September 2010.
15. Sogin S, Lai YC, Dick CT, et al. Comparison of capacity of single- and double-track rail lines. *Transp Res Rec J Transp Res Board* 2013; 2374: 111–118.
16. Sogin S, Lai YC, Dick CT, et al. Analyzing the transition from single- to double-track railway lines with nonlinear regression analysis. *J Rail Rapid Transit* 2016; 230: 1877–1889.
17. Boysen HE. General model of railway transportation capacity. *Computers in Railways XIII, WIT Transactions on the Built Environment* 2012; 127: 335–346.
18. Boysen HE. Quicker meets, heavier loads and faster empties – effects on transportation capacity and cycle time. In: *Proceedings of 10th International Heavy Haul Association*, New Delhi, India, 4–6 February 2013, pp.838–844. International Heavy Haul Association (IHHA).
19. Lindfeldt O. *Railway operation analysis: evaluation of quality, infrastructure, and timetable on single and double-track lines with analytical models and simulation*. PhD Thesis, KTH Royal Institute of Technology, Department of Transport, Sweden, 2010.
20. Goverde RMP, Corman F and D'Ariano A. Railway line capacity consumption of different railway signaling systems under scheduled and disturbed conditions. *J Rail Transp Plann Manage* 2013; 3: 78–94.
21. Dick CT and Mussanov D. Operational schedule flexibility and infrastructure investment: capacity trade-off on single-track railways. *Transp Res Rec J Transp Res Board* 2016; 2546: 1–8.
22. Sehitoglu T, Mussanov D and Dick CT. Operational schedule flexibility, train velocity and the performance reliability of single-track railways. In: *Proceedings of the transportation research board 97th annual conference*, Washington, DC, United States, 7–11 January 2017.
23. Mussanov D, Nishio N and Dick CT. Delay performance of different train types under combinations of structured and flexible operations on single-track railway lines in North America. In: *Proceedings of the International Association of Railway Operations Research (IAROR)*, Lille, France, 4–7 April 2017.
24. Berkeley Simulation. Rail Traffic Controller (RTC) user manual, July 2011.
25. Mussanov D, Nishio N and Dick CT. Building capacity through structured heavy haul operations on single-track shared corridors in North America. In: *Proceedings of 11th International Heavy Haul Association*, Cape Town, South Africa, September 2017.