

Analyzing the transition from single- to double-track railway lines with nonlinear regression analysis

Proc IMechE Part F:
J Rail and Rapid Transit
2016, Vol. 230(8) 1877–1889
© IMechE 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0954409715616998
pif.sagepub.com



Samuel L Sogin¹, Yung-Cheng (Rex) Lai², C Tyler Dick¹
and Christopher PL Barkan²

Abstract

The demand for freight rail transportation in North America is anticipated to substantially increase in the foreseeable future. Additionally, government agencies seek to increase the speed and frequency of passenger trains operating on certain freight lines, further adding to demand for new railway capacity. The majority of the North American mainline railway network is single track with passing sidings for meets and passes. Expanding the infrastructure by constructing additional track is necessary to maintain network fluidity under increased rail traffic. The additional track can be constructed in phases over time, resulting in hybrid track configurations during the transition from purely single track to a double-track route. To plan this phased approach, there is a need to understand the incremental capacity benefit as a single-track route transitions to a two-main-track route in the context of shared passenger and freight train operations. Consequently, in this study, the Rail Traffic Controller software is used to simulate various hybrid track configurations. The simulations consider different operating conditions to capture the interaction between traffic volume, traffic composition and speed differences between train types. A nonlinear regression model is then developed to quantify the incremental capacity benefit of double-track construction through exponential delay–volume relationships. Adding sections of double track reduces train delay linearly under constant volume. This linear delay reduction yields a convex increase in capacity as double track is installed. These results allow railway practitioners to make more-informed decisions on the optimal strategy for incremental railway capacity upgrades.

Keywords

Rail transportation, simulation, shared corridor, regression analysis

Date received: 8 June 2015; accepted: 16 October 2015

Introduction

Most of the North American railway network is single track with passing sidings (also known as passing loops) for meets and passes between trains. Of all mainlines with freight traffic exceeding 10,000,000 gross tons per year, only 37% have multiple tracks.¹ Future demand for increased freight and passenger rail service will require more railway network capacity. Consequently, a considerable number of single-track routes will require additional tracks to accommodate traffic demand. There are three basic approaches to increasing the capacity of a single-track line: extending the length of passing sidings, adding passing sidings, and adding double track. Extending passing sidings enables longer freight trains and reduces passenger delays from meets with other trains. Additional passing sidings are typically installed in the longest single-track bottleneck between existing passing sidings. After implementing these intermediate solutions, double track may

become the most-effective way to meet additional traffic demand. The second mainline track can be constructed in phases over time, such that the amount installed matches expected increases in rail traffic.² In this paper, we refer to intermediate phases with partial double track as “hybrid” track configurations. Most railway infrastructure in the United States is owned by private freight railway companies.³ With few exceptions, public passenger train operating agencies must negotiate access to freight railway lines to provide passenger service. In most circumstances, the freight railway requires that its level of service not be

¹Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, USA

²Department of Civil Engineering, National Taiwan University, Taiwan

Corresponding author:

Yung-Cheng (Rex) Lai, Room 313, Civil Engineering Building No. 1
National Taiwan University, Roosevelt Road, section 4, Taipei 10617,
Taiwan.

Email: yclai@ntu.edu.tw

negatively impacted by the passenger traffic added to the route. To ensure performance is maintained and mitigate potential delays, additional track is usually installed. In some situations, the passenger agencies may pay freight railroads for the slots their passenger trains consume and the required track expansion.

Past research and practical experience have demonstrated that single- and double-track lines perform in a very different way.^{4,5} Single track has considerably lower capacity than double track. The key reason behind this distinction in capacity is that trains running in opposite directions take turns using the single-track sections between passing sidings. These single-track sections are often bottlenecks that constrain overall line capacity. On double-track mainlines, since one track can essentially be dedicated to a particular direction of traffic, capacity is primarily affected by the minimum spacing and headway between subsequent trains traveling in the same direction. Capacity of double track is reduced under heterogeneous traffic due to possible overtakes, however, even in these conditions it is still much higher than the capacity of single track.⁶

Existing literature on single-track operations has shown that adding a high-priority train, such as a passenger train, to a freight network significantly increases the average train delay compared with adding a train with similar characteristics to others operating on the line. Vromans et al.⁷ used simulation to investigate strategies to improve passenger operations. Leilich⁸ and Dingler et al.⁹ used simulation to evaluate the interactions between bulk freight trains and high-priority container trains, and found a capacity loss due to the heterogeneous operations. Sogin et al.¹⁰ simulated single-track shared corridors and determined that changes in the maximum speed of the high-priority train had little relationship with train delay, as the primary delay comes from meets at sidings. However, double-track operations were found to be more sensitive to the speed differential among different types of trains. The second main track may be used by the high-priority train to overtake slower trains, resulting in an exponential delay distribution where many trains operate close to the minimum run time.¹¹

There has been limited research on hybrid track configurations. Petersen and Taylor¹² used simulation to locate longer passing sidings in order to accommodate passenger trains on a freight line. Additionally, Pawar¹³ used analytical models to determine the length of long passing sidings required to run a single-track, high-speed railway without meet delays. Lindfeldt² compared partial double track to additional passing sidings and determined that partial double track offers more timetable flexibility and improves practical capacity relative to additional passing sidings.

This research focuses on capacity benefit as a single-track route transitions to a two-main-track

route in the context of shared passenger and freight train operation. Railway traffic simulation software was used to evaluate each intermediate phase of double-track installation at different traffic levels. A typical North American single-track route was used as the baseline condition. Sections of double track were systematically added to the baseline route by connecting pairs of pre-existing passing sidings. Train delay and capacity transition curves were then described mathematically. This procedure was first applied to a homogeneous freight corridor. It was then compared to a shared corridor where 25% of the total traffic was passenger trains. In order to differentiate between delay mechanisms, the shared corridor analysis was conducted twice. Initially, to determine the impact of priority, the speed limit of the passenger trains was the same as the freight trains, but with higher priority. The second analysis determined the marginal effect of a speed differential by using higher-speed, higher-priority passenger trains. Through this analysis, the capacity impact of the passenger trains on freight railway operations can be attributed to specific delay-causing mechanisms. These results can aid railway planners in determining the amount of double track required to mitigate the effect of additional traffic on a rail corridor. The results of these experiments provide a better understanding of key fundamental relationships affecting railway performance.

Delay as a Proxy for Capacity

Measuring railway capacity is a non-trivial problem. Theoretical capacity can be measured using analytical techniques; however, when measuring practical capacity it can be difficult to incorporate all of the stochastic factors affecting train operations. In the United States, it is common practice to simulate current railway traffic on the existing infrastructure and then re-execute the simulation with additional traffic. The differences in delay between these two cases are analyzed and usually train delays increase. The incremental delay can be mitigated by constructing additional track. A series of alternative infrastructure configurations is then simulated, and generally the one yielding the best return on investment is selected for construction. This process does not explicitly calculate the practical capacity of the rail line.

Several studies have developed regression models to quantify the impact of various operational factors on train delay.¹⁴⁻¹⁷ Another approach to determining railway capacity, is to define a mathematical relationship between traffic volume and train delay through regression techniques. Using this method of delay-volume curves, train delay can be predicted as an exponential function of traffic volume^{15,17,18} as

$$D = Ae^{kV} \quad (1)$$

where D is the delay mitigation constant, A is the congestion factor, k is the average train delay and V is the traffic volume.

The relationship between train delay and traffic volume is characterized by the two key coefficients in the exponential function, i.e. the delay mitigation constant (A) and the congestion factor (k). The delay mitigation constant (A) reflects the ability of the route to absorb delays given its infrastructure, traffic and operating characteristics. The congestion factor (k) captures the sensitivity of delay and level of congestion to added traffic volume. Both values are unique to a specific segment of a corridor; based on its infrastructure, traffic and operating conditions.^{15,17} According to Krueger¹⁵, a typical value of A is 0.2–0.6, and a typical value of k is approximately 0.048.

Using this relationship, a railway could define the capacity of a line as the traffic volume (number of trains per day) where the level of service deteriorates to a minimum level of service (MLOS) that is still acceptable. The exact definition of MLOS differs depending on the infrastructure owner and railway operator. In this research, MLOS is defined as the maximum average train delay that is tolerable to the railway operator, D_{max} . Under this definition of MLOS, equation (1) can then be rearranged and solved for railway capacity

$$V = \frac{1}{k} \ln\left(\frac{D_{max}}{A}\right) \tag{2}$$

Consider two different single-track routes that have differing amounts of double-track sections installed. The delays on these two routes can be characterized by equation (1). Assuming that each route is operating at the traffic volume corresponding to the MLOS,

D_{max} , then using equation (3a), the difference in capacities can be solved. The capacity difference of the capacities of these two lines is independent of the specified MLOS

$$V_2 - V_1 = \frac{1}{k_2} \ln\left(\frac{D_{max}}{A_2}\right) - \frac{1}{k_1} \ln\left(\frac{D_{max}}{A_1}\right) \tag{3a}$$

$$V_2 - V_1 = \left(\frac{1}{k_1} + \Delta\right) \ln\left(\frac{D_{max}}{A_2}\right) - \frac{1}{k_1} \ln\left(\frac{D_{max}}{A_1}\right) \tag{3b}$$

$$V_2 - V_1 = \frac{1}{k_1} \ln\left(\frac{A_1}{A_2}\right) + \Delta V_2 \tag{3c}$$

If the congestion factors for the two routes are approximately equal then Δ would be approximately equal to zero. In this case, equation (3c) can be approximated by equation (4). This may be a reasonable assumption when the types of traffic interactions on two infrastructure configurations are similar. However, a homogenous freight line may have a congestion factor that changes under mixed passenger and freight traffic. The change in capacity described by equation (4) is independent of any delay standard that a railway might set. Equation (4) could be used as a base for comparing the capacity improvement by adding sections of double track if the three coefficients can be related to the amount of double track installed

$$V_2 - V_1 = \frac{1}{k} \ln\left(\frac{A_1}{A_2}\right) \tag{4}$$

We presume that there is some functional relationship describing the relationship between capacity and the percentage of double track installed. Figure 1 depicts five possible transition functions from single

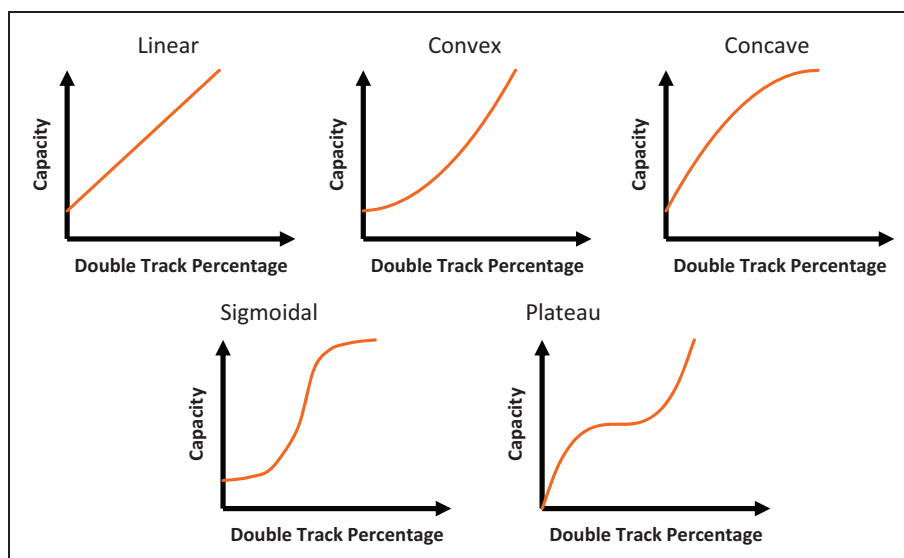


Figure 1. Potential shapes of transition functions from single to double track.

to double track. These five curves are all upward-sloping assuming that the capacity increases as additional track is added. If capacity were measured on the *y*-axis, then there would be an assumed positive relationship between capacity and the double-track percentage. In this paper, we aim to identify the functional relationship for train delay and capacity under various transition scenarios from single to double track. The shapes of these curves may differ for different performance metrics.

Simulation Methodology

The hybrid track configuration experiment examined traffic volume, traffic mixture and the percentage of the second mainline track installed; and used train delay as the response variable. The original single-track line parameters are summarized in Table 1. This baseline is typical of a high-quality, single-track mainline in North America with high turnout entry speeds, dense passing siding spacing, and capacity for long trains.

Different strategies regarding the phased construction of second mainline track can be implemented on the corridor. Figure 2 demonstrates two possible strategies for determining where to construct additional tracks along the route: *alternate* and *split*.

The *alternate* strategy is to choose several (e.g. four to six) locations on the line and build the additional track in both directions from these intermediate points. This has the benefit of creating long sections of double track where two trains can meet without either train having to stop at passing sidings. Additionally, these double-track sections may be long enough to achieve an overtake maneuver, where a faster train uses one track to overtake a

slower train running or stopping on the other track. Another possible strategy is to *split* the double-track resource and then build the additional track from the two terminals towards the midpoint. The *split* strategy has the advantage of easing potential bottlenecks at or near terminals. This provides longer double-track sections than the alternate condition, with the trade-off being a longer section of single track in the middle of the route between the double-track sections at the ends.

In both allocation strategies, double track is added to connect pairs of pre-existing passing sidings, linking them together to form the second mainline track. The double track is installed to minimize reconfiguration of turnouts and signal system control points. When a section of double track connects a passing siding as mainline track, the turnout for that passing siding is reused as part of a future universal crossover leading into a future section of double track, resulting in crossovers spaced approximately 10-miles apart on the double-track lines. When a passing siding becomes part of a section of second mainline track, its track speed is upgraded to match the rest of the mainline. All tracks in the simulation are able to handle bi-directional traffic, as is common North American railroad practice. Although flexible operation raises the possibility for overtakes, this kind of movement rarely happens, due to it being more efficient to run directional traffic in the double-track network, especially at high traffic levels. The amount of second track installed is described by the double-track percentage. For the purposes of this paper, the double-track percentage includes both the length of passing sidings and the second mainline track. In this case, the baseline single-track configuration had 45.6 miles of passing sidings and is therefore classified as 19% double track.

Rail traffic controller (RTC) software, developed by Berkeley Simulation Software, is the standard tool for railway simulation analysis in North America.¹⁹ Users include all U.S. Class I railways, Amtrak, the Surface Transportation Board, Bay Area Rapid Transit, major consulting firms and many others. RTC calculates train movements over a route while considering allowable track speeds, grade, curvature, and signal systems. RTC can modify train paths when trains conflict with each other, such as two trains simultaneously requesting use of the same section of track. Once a conflict is recognized, the simulation logic acts as a dispatcher to delay and/or reroute trains based on their priorities. Train priorities and departure times are initially specified by the users. As the simulation proceeds, the priority of a train varies as a function of its on-time performance. For example, late trains are given priority over early trains. Additionally, the priorities can be adapted to reflect business objectives, such as giving preference to intermodal container trains over bulk commodity trains, or passenger trains over

Table 1. Route parameter guidelines.

Parameter	Value
Length	240 miles
Bottleneck length	8 miles
Siding spacing (on center)	10 miles
Siding length	2 miles
Diverging turnout speed	45 mile/h
Traffic control system	2-Block, 3-Aspect CTC
Average signal spacing	2 miles

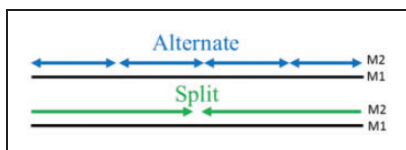


Figure 2. Double-track allocation strategies.

freight trains. The architecture behind RTC is shown in Figure 3.

In this study, for a specific traffic and infrastructure scenario, an RTC simulation run examines 5 days of operations. The simulation is repeated six times to generate performance statistics for 30 days of operations. Replication allowed the dispatching algorithm to make different decisions with similar inputs to the model.

Although the infrastructure differs among cases in order to represent the varying amount of double track, the boundary conditions are kept constant. The route features only one origin–destination pair and traffic is directionally balanced. Each end of the route features terminals designed to minimize terminal–mainline interference by having long leads, and multiple receiving and departure tracks.

In addition to the infrastructure and operating characteristics, scheduling also affects the operational performance of a corridor. In many railway operations, conflicts between trains are carefully planned in a timetable. However, railway operations in North America are highly variable due to fluctuations in freight traffic demands, weather, length of haul, and

other sources of variation and delay. Train dispatchers resolve conflicts between trains in real-time instead of following a strict timetable. Different train schedules on the same infrastructure can show different average train delays. Consequently, assuming only one schedule for a particular infrastructure may incur large experimental error. To address this issue, the departure time of each train over a 24-h period is determined from a random uniform distribution for each simulation trial. In this way, stable averages can be obtained by averaging the operational performance over varying train schedules over multiple days.

Table 2 demonstrates the characteristics of the passenger and freight trains in the simulation. The freight train characteristics are based on several past studies on railway capacity.^{20,21} The characteristics of passenger train represent the 110-mile/h passenger trains operating between Chicago and Detroit. The passenger train stops at approximately 32.4-mile intervals based on the average station spacing on regional intercity passenger rail routes in California, Illinois, Washington, and Wisconsin.²²

In this study, train delay is selected as a proxy for capacity, as large delays indicate a congested network.

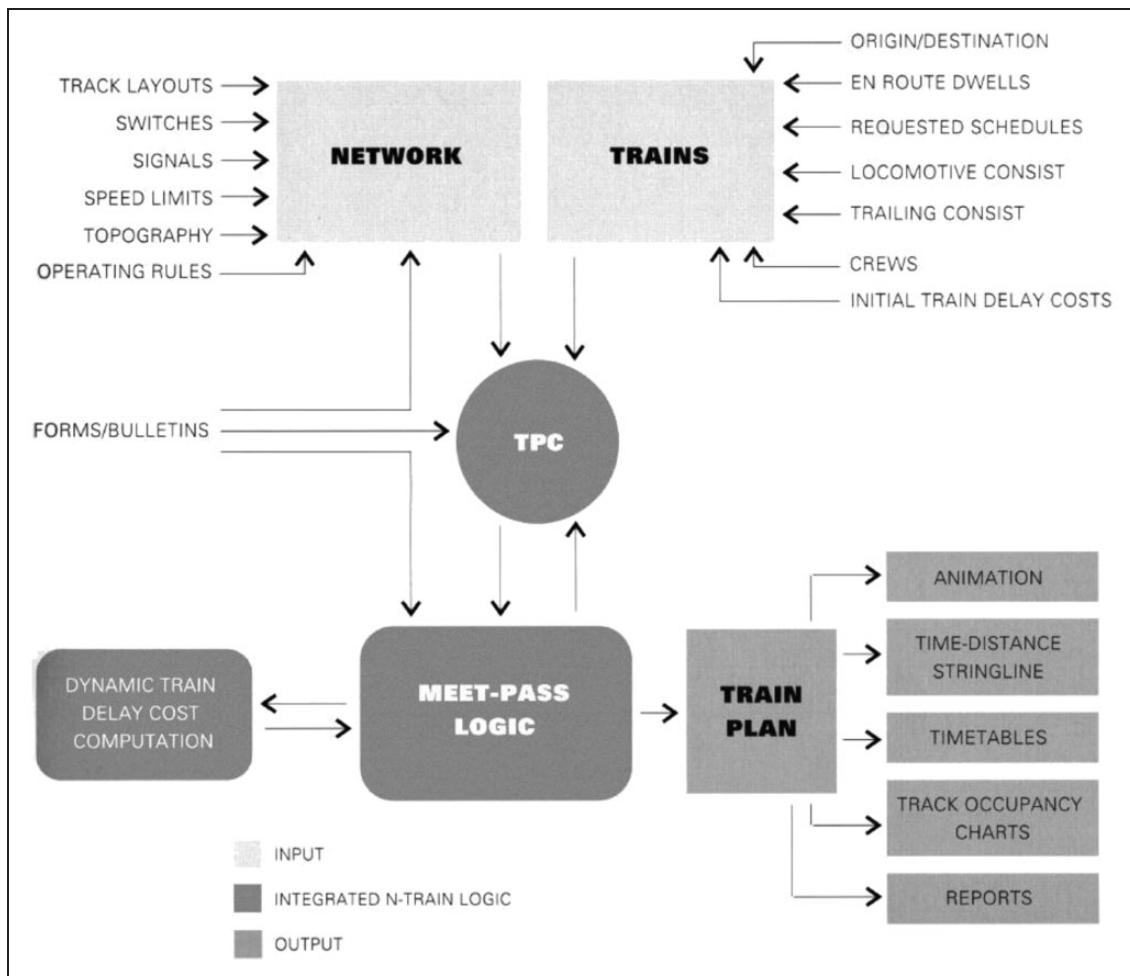


Figure 3. Architecture of RTC.¹⁹

Passenger traffic has a lower tolerance for delay compared with most freight traffic. These analyses focus on delays to freight trains, as these delays are most responsive to the identified factors. This is due to the passenger trains being shielded from high delays by their high priority.

Previous work on shared corridors has identified two major characteristics of passenger trains that may cause delays to freight trains: higher priority and speed differentials.^{10,11} The effect of priority was analyzed by having a high-priority passenger train travel at the same speed as the freight trains (50 mile/h). The effect of speed differential and priority acting together was represented by 110-mile/h high-priority passenger trains operating with 50-mile/h low-priority freight trains. In the context of the simulation software, priority is a measure of preference. There may be situations where delaying a passenger train can result in better network fluidity.

For example, an eastbound passenger train may stop on a passing siding to meet two consecutive oncoming westbound freight trains. This can result in lower network delays than by splitting this conflict into two separate meet conflicts at two passing sidings. In single-track networks, passenger trains are often delayed by meets with other high-priority passenger trains traveling in opposite directions, as the siding length dictates that at least one train stops.

Developing the Response Surface Model

The following analysis focuses on developing a response surface model based on the simulation data. The goal is to predict the capacity of a rail line as a function of the amount of double track installed and the MLOS. The analysis in this section shows the development of a response model for a freight-only corridor where the double track is allocated in an *alternate* strategy.

The evidence from this study suggests that train delay decreases linearly for each marginal section of double track added to the single-track baseline condition using an *alternate* strategy. This linear decrease in train delay occurs in each of the eight different traffic levels studied from eight trains per day (TPD) to 64 TPD. Figure 4 shows the freight train delays at the eight different traffic levels over 14 different track configurations progressing from pure single track (19%) to complete two-mainline track (100%). The linear reduction in train delay is greater with higher traffic levels than with lower traffic levels. Additionally, these trend lines converge around approximately 100% double track.

Table 2. Characteristics of freight and passenger trains.

	Passenger train	Freight train
Locomotives	x2 P42	x3 SD70
Number of railcars	8 single-level cars	115 hopper cars
Length (ft)	740	6325
Weight (tons)	800	16,445
Horsepower per ton	15.4	0.78
	32.4 miles between stops	

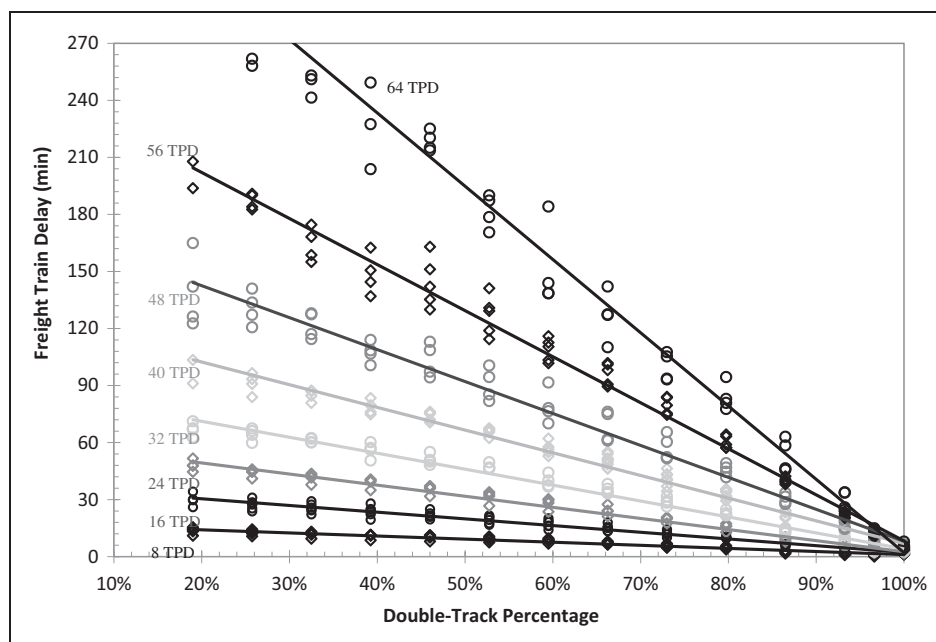


Figure 4. Train delays as a function of percentage of double track at various traffic volumes.

Each of the trend lines in Figure 4 can be described by its slope and intercept parameters. If there are clear relationships between these parameters and traffic volume, there may be a master equation that predicts delay on the route for a given double-track percentage and traffic volume. This equation would be in the form of equation (5). The intercept $\gamma_0(V)$ and slope $\gamma_1(V)$ are both functions of volume. In an alternate formulation, the slope term could represent the delay reduction per mile of double track installed under constant volume. In this analysis, the double-track parameter is normalized by route length, so the slope parameter is the reduction in train delay per double-track percentage point. An important property of equation (5) is that it is centered on 19%, the single track configuration. This point-slope format results in the intercept term that relates to the amount of train delay in a single-track configuration. Otherwise, in slope-intercept format, the y-axis

intercept would indicate a theoretical amount of delay on a single track with zero passing sidings. Values in this range were not simulated and violate the route parameter guidelines in Table 1. The x-axis intercept is an indicator of the level of double-tracking where the line experiences no train delays. In the cases simulated, this value was greater than 100% and indicates a small amount of triple track

$$D(V, x) = \gamma_0(V) - \gamma_1(V) \times (x - 19\%) \tag{5}$$

where $\gamma_0(V)$ is the single-track train delay as a function of traffic volume (intercept), $\gamma_1(V)$ is the reduction in train delay per double-track percentage point as a function of traffic volume (slope), and x is the double-track percentage

The parameters of the linear trend lines for each traffic volume are shown in Table 3. The single-track intercepts are always positive and increase with higher traffic volumes. The negative slope terms increase in magnitude with higher traffic volumes (Figure 4). These trends in opposite directions can describe the pivoting of the trend lines around an approximately full double track in Figure 4. Both the slope and single-track-delay intercept parameter are plotted as points against volume in Figure 5. The relationship between these trend-line parameters and traffic volume can be explained by several different functional relationships, including exponential or polynomial. An exponential relationship only requires two parameters to describe $\gamma(V)$, whereas a polynomial would require at least three, assuming at least a quadratic fit. In higher-order polynomials, there is significant difficulty in deriving physical meaning from parameter estimates. Lastly, using an exponential relationship is more likely to simplify to an

Table 3. Linear parameter estimates of train delay under constant volume.

Volume (TPD)	Slope, $\gamma_1(V)$ (min/double track %)	Single-track delay, $\gamma_0(V)$ (min)
8	-16.3	14.2
16	-35.0	30.6
24	-58.4	49.8
32	-83.9	72.0
40	-117.1	102.3
48	-168.1	144.2
56	-239.9	203.1
64	-385.5	314.4

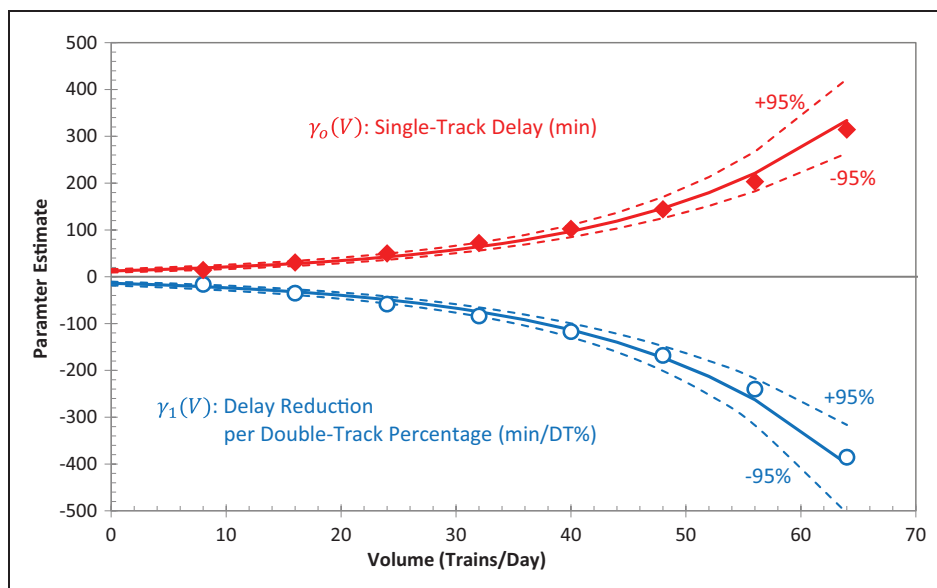


Figure 5. Linear parameter estimates of train delay reductions by adding sections of two-mainline track. These parameter estimates are then predicted by exponential relationships.

equation similar to equation (1). If exponential relationships are assumed then equation (5) becomes equation (6)

$$D(V, x) = A_0 e^{k_0 V} - A_1 e^{k_1 V} \times (x - 19\%) \quad (6)$$

where k_2 is the slope congestion factor.

The solid lines in Figure 5 predict the linear parameters using exponential relationships. The single track-intercept parameter can be predicted using equation 7(a) and the slope by equation 7(b). These relationships were determined using a log-transformation and simple linear regression procedures in JMP.²³ The dashed lines represent 95% confidence bands around the mean response of the linear parameter estimate. Equations 7(a) and 7(b) can accurately predict the linear parameter estimates and be substituted into equation (6) to yield equation (8)

$$\gamma_0(V) = 12.27 e^{0.05162V} \quad (7a)$$

$$\gamma_1(V) = 13.89 e^{0.05249V} \quad (7b)$$

$$D(V, x) = 12.269 e^{0.05162V} - (13.889 e^{0.05249V})(x - 19\%) \quad (8)$$

A disadvantage in the method of producing equation (8) is that each time the data are passed through a simple linear regression, degrees of freedom are lost. In this case, eight linear trend lines were determined, each featuring two parameter estimates. Equations 7(a) and 7(b) also require two parameter estimates. In total, 20 different parameter estimates were determined to derive equation (8). As an alternative to this hierarchical regression approach, nonlinear regression can be used to arrive at the final four parameter estimates of equation (6) without the same loss in degrees of freedom. Not surprisingly, the nonlinear regression platform yields more precise parameter estimates and results in a lower root mean square error (RMSE) of the original data.

An important aspect of the parameter estimates for equation (6) in Table 4 is the similarity between A_0 and A_1 as well as the estimates for k_0 and k_1 . With 95% confidence intervals, there is clear overlap between these pairs of parameter estimates. Although the form of equation (8) is significant and built on sound theory, a simpler model may be sufficient. In particular, if k_0 and k_1 are equal, then the freight train delay data can be described by equation (9). This equation is no longer centered on 19% double track for the purposes of simplicity. Equation (9) is in the form of equation (1) where the A term is now described by a linear function of the double-track percentage. An interesting property of equation (9) is that it has a closed-form solution when solved for traffic volume instead of delay as shown in equation (10). In this form, the capacity of

Table 4. Comparison of the hierarchical and nonlinear regressions.

Statistic	Hierarchical	Nonlinear
RMSE	31.04	8.29
Single track delay intercept		
A_0		
Estimate	12.269	13.201
Lower limit	9.237	12.554
Upper limit	16.298	13.881
k_0		
Estimate	0.05162	0.0495
Lower limit	0.04459	0.0486
Upper limit	0.05864	0.0504
Slope (delay reduction)		
A_1		
Estimate	13.889	14.345
Lower limit	10.478	13.197
Upper limit	18.410	15.594
k_1		
Estimate	0.05249	0.0513
Lower limit	0.04551	0.0498
Upper limit	0.05947	0.0527

the line is then a function of the amount of double track installed and a delay standard, D_{\max}

$$D = (S_1 - S_2 x) e^{kV} \quad (9)$$

where S_1 is the single-track delay constant, S_2 is the delay mitigation constant and k is the congestion factor

$$V = \frac{1}{k} \times \ln\left(\frac{D_{\max}}{S_1 - S_2 x}\right) \quad (10)$$

Equation (10) is plotted in Figure 6 for different delay standards. The capacity improvement of the double-track installation is close to linear when the route is closer to a single-track configuration. As more double track is added and the route approaches a full two-main-track configuration, the additional segments of second track yield increasingly greater capacity benefits. These capacity curves can help justify last-mile investments to complete the double track on a line. These final projects may correspond to expensive tunnels, bridges, mountain passes or improvements in urban areas. The delay standard has more of an effect on determining the capacity of the line when the standard is low. At higher delay standards, the capacity contours are grouped much more closely (Figure 6).

The instantaneous slopes of the curves plotted in Figure 6 are parallel at all double-track percentages. This property is verified by taking the partial

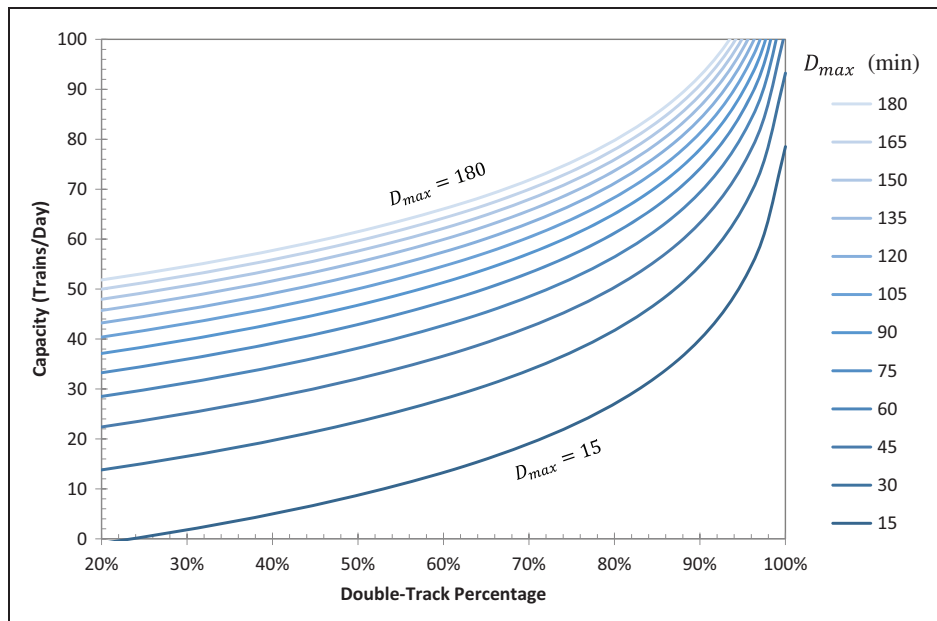


Figure 6. Capacity as a function of the amount of double track installed under different delay standards D_{max} .

derivative of equation (10) with respect to the double-track percentage (equation (11)). The implication is that the change in capacity from installing sections of double track is independent of the delay standard D_{max} . For example, capacity increases by 5.8 TPD at each delay standard when upgrading from 60% double track to 70% double track. As equation (11) does not include the delay standard, D_{max} , changing the delay standard would result in a constant change in capacity that is independent of the double-track delay percentage, x . For example, changing the delay standard from 30 to 60 min would increase capacity by 14.2 TPD regardless of double-track percentage

$$\frac{\partial V}{\partial x} = \frac{S_2}{k(S_1 - S_2x)} \quad (11)$$

Comparing different operating conditions

The previous analyses were completed using only homogenous freight trains and only one method of allocating sections of mainline track to a railway corridor. In the following analyses, the effect of additional parameters was considered using equations (9) and (10). First, the double-track allocation strategy is changed from *alternate* to *split* (Figure 2). Instead of simulating eight different traffic levels for all infrastructure configurations, only traffic levels of 16, 40 and 56 TPD were considered in order to develop appropriate parameter estimates for equation (9). The capacity of two or more configurations is compared by using equation (10) at a MLOS set to a 60-min average delay. In the previous section, the

change in capacity was independent of the MLOS. In order to use equation (10) to compare different operating conditions, six different parameters must be estimated and the MLOS does not drop out of the model. Fortunately, the partial derivative of volume with respect to double-track percentage is much greater than the partial derivative of volume with respect to D_{max} , (equation (12)). In the case of the freight-only corridor at 60% double track and a 60-min MLOS, using an alternate strategy, a unit change in double track is greater than a unit change in the delay standard by a factor of 143. As long as the parameter estimates of equation (9) are of the same magnitude as those estimated in the previous section, then the MLOS has a small effect on the change in capacity between different operating conditions

$$\frac{\partial V}{\partial D_{max}} = k(D_{max})^{-1} \quad (12)$$

The two allocation strategies for homogenous freight trains do not lead to significantly different changes in line capacity. Allocating track in a *split* configuration instead of an *alternate* configuration shows an improvement of about one-half TPD (Figure 7). The parameters for these two scenarios are summarized in Table 5. Although there may be little change between the two infrastructure configurations under the operating conditions of this analysis, there may be greater changes in the results under different scenarios. For example, if more sophisticated terminal effects were included in the model then the split configuration may show more benefit. If the model followed a strict timetable, then meets between trains can be planned to occur at sections of double

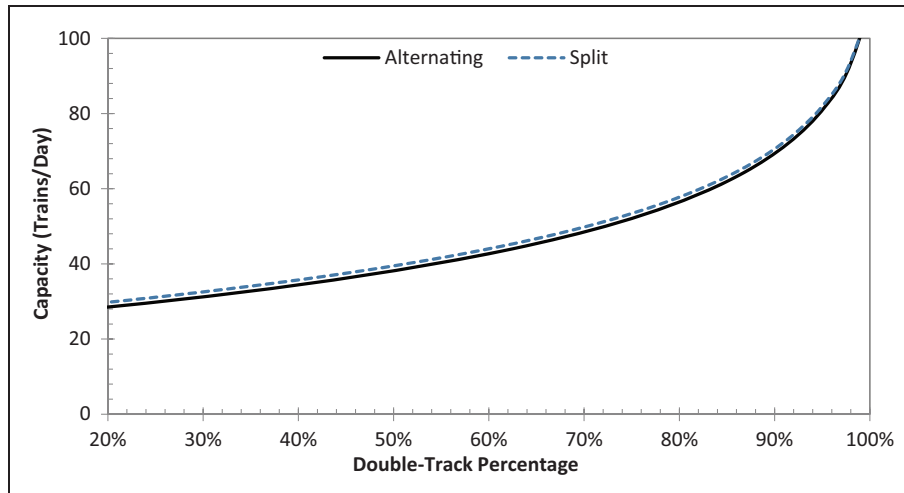


Figure 7. Capacity improvement under a 60-min delay standard when sections of double track are allocated to the terminals (split) or at a collection of midpoints along the line (alternate).

Table 5. Parameter estimates for different allocation strategies.

	Alternate	Split	Percentage change (%)
S_1	19.5206	18.4404	-5.53
S_2	19.1490	18.0585	-5.69
k	0.0471	0.0469	-0.42

track and take better advantage of the *alternate* configuration.

We considered two potential mechanisms that might cause additional freight train delays due to passenger trains on lines where they share tracks. The first is different priorities between train types where passenger trains are given preference in meet or pass conflicts. The second delay mechanism studied was differences in speed between train types. The effect of priority was illustrated on a mixed-traffic line where there are three freight trains per passenger train. Both trains were limited to a maximum speed of 50 mile/h. The effect of speed and priority were evaluated by having 25% of the total traffic comprised of 110-mile/h passenger trains and by having 50% of the total traffic comprised of 79-mile/h passenger trains (most common speed limit for passenger trains without a cab signaling system). The double track was allocated using the *alternate* strategy.

The general trend for freight only (0% passenger trains), 25% 50-mile/h passenger trains and 50% 79-mile/h passenger trains are quite similar. A heterogeneous mixture of three freight trains per passenger train would by itself result in a capacity loss for any delay standard. Higher heterogeneity and more passenger trains further reduce the capacity level due to the operational constraints from different types of trains. Additionally, it takes more double track to

Table 6. Parameter estimates of equation (9) for different traffic types.

	Freight only	50-mile/h passenger trains	110-mile/h passenger trains	79-mile/h passenger trains
S_1	19.5206	19.9317	22.4534	19.9595
S_2	19.149	19.3509	20.4052	19.5561
k	0.0471	0.0547	0.0495	0.07316

mitigate traffic increases. Having priority trains only manifests a change relative to the base case by having a higher k coefficient, indicating a higher sensitivity to traffic increases. The higher speed passenger trains have a k coefficient on par with the freight-only case; however, they also have higher S_1 and S_2 coefficients (Table 6). The capacity curves for 60-min MLOS are plotted in Figure 8. The difference between 50-mile/h and 110-mile/h passenger trains is very small, for between 20 and 80% double track. In the 80 to 100% double-track range, there is a divergence between the two passenger train interference curves, where speed differentials start to reduce capacity.

The loss in capacity due to 110-mile/h passenger trains in a freight corridor under an average 60-min MLOS is illustrated in Figure 9. The capacity loss due to the higher-speed passenger train is greatest on double track and lowest on single track. Additionally, this loss curve has a convex transition where the change in capacity is greater at higher double-track percentages than at lower percentages. The dashed curve of Figure 9 estimates the delay mechanism by comparing the capacity loss due to priority with 50-mile/h passenger trains, to the total loss from 110-mile/h passenger train interference. The effect of the priority differential accounts for 96% of the capacity loss between 19 and 55% double track.

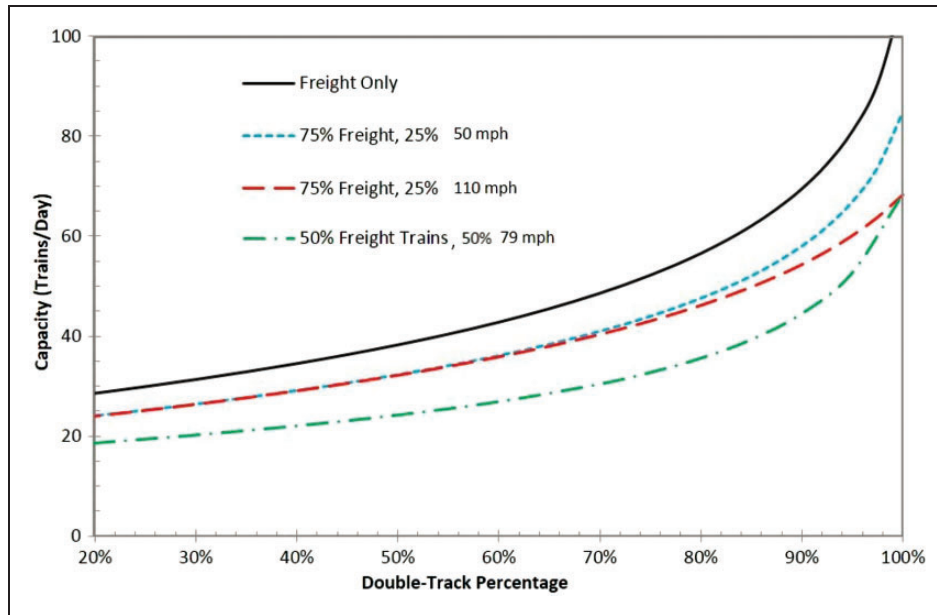


Figure 8. Change in capacity by installing sections of double track with a 60-min delay standard for traffic configurations of: 100% freight trains; 75% 50-mile/h freight trains and 25% 50-mile/h passenger trains; 75% 50-mile/h freight trains and 25% 110-mile/h passenger trains; and 50% 50-mile/h freight trains and 50% 79-mile/h passenger trains.

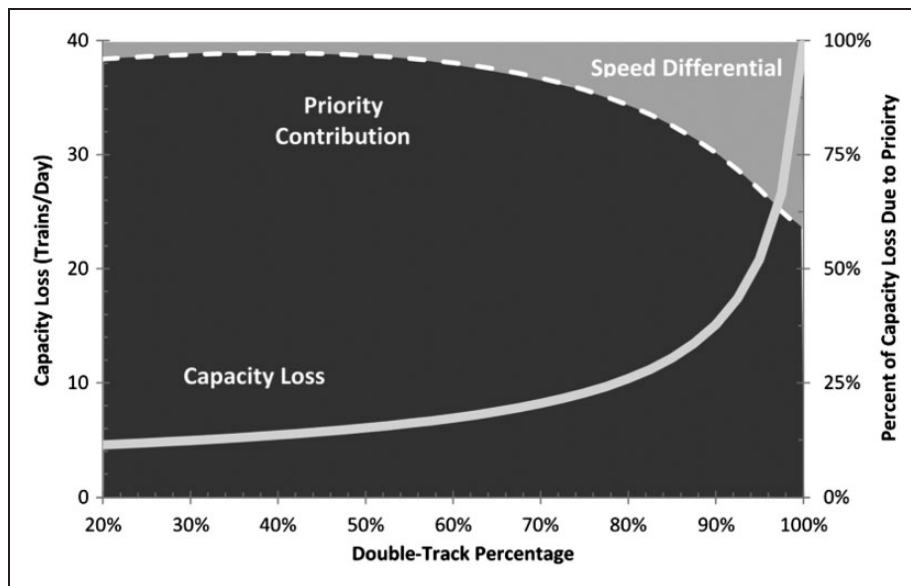


Figure 9. Capacity lost due to the operation of higher-speed passenger trains and the estimated contribution due to the delay mechanism.

At full 100% double track, the speed differential mechanism accounts for 41.7% of the loss in capacity, and priority accounts for 58.3%.

Heterogeneity in priority and the speed differential has a significant impact on capacity. The freight train capacity loss on a higher-speed passenger, mixed-use corridor is even greater than that shown in Figure 9, due to it assuming that 25% of the available capacity is being used to accommodate passenger trains instead of freight trains. Consider a case where a railway line is originally a single track, freight-only corridor

at full capacity. The long-term plan is to change this line into a mixed-use corridor where future traffic is comprised of 75% freight and 25% 110-mile/h passenger trains. The initial capital investment mitigates the additional delays to the original freight trains and improves capacity by 33% to accommodate the additional passenger trains. If the freight line was originally single track, then this initial investment would be to upgrade the line from 19% double track to 65% double track under a 60-min MLOS.

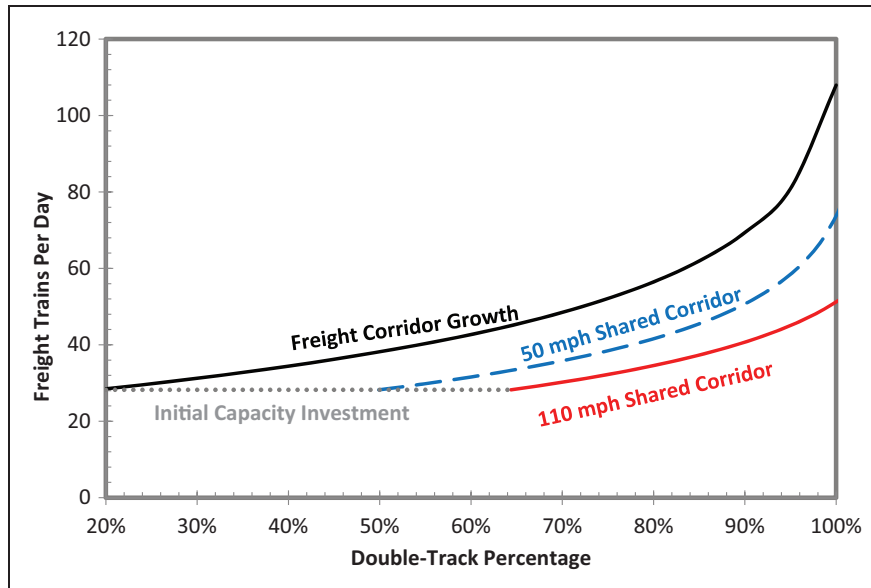


Figure 10. Freight train capacity grown under a 60-min MLOS for a freight corridor and a shared corridor comprised of 75% freight and 25% passenger trains.

The amount of double track needed to accommodate the passenger trains was calculated using equation (13)

$$x_m = S_{2_p}^{-1} \left(S_{1_p} - (D_{\max}) e^{-V_0 k_p \varphi^{-1}} \right) \quad (13)$$

where x_m is the level of double tracking to mitigate addition of passenger service, S_{1_p} is the shared corridor single-track delay constant, S_{2_p} is the shared corridor delay mitigation constant, k_p is the shared corridor congestion factor, V_0 is the initial freight corridor volume (capacity at D_{\max}) and φ is the freight train percentage of the total traffic of the planned shared corridor.

This substantial investment in capacity to accommodate passenger trains does not benefit the freight railway; it simply allows them to maintain their current level of service. In this 75% freight, mixed-use corridor, any future growth in both freight and passenger volume would require additional capacity investment with 65% double track as the new baseline (Figure 10). If there were significant engineering cost constraints to expanding this corridor beyond two-main tracks (i.e. triple track), then the freight railway has lost its ability to accommodate new freight business in the future. If the speed differential between different types of trains were eliminated, then the initial capacity investment would be to upgrade the corridor to 50% double track. With 50-mile/h passenger trains, there is also more freight capacity available in a full two-track scenario than with 110-mile/h passenger trains.

Fortunately for the freight railway, there are some short-term benefits. By having a passenger agency

make the initial investments to accommodate the passenger trains, the next time freight demand increases, the freight railway receives a higher return on capacity per track-mile installed. This benefit occurs because the freight railway is now on the 110-mile/h shared corridor curve in Figure 10, which is steeper at 65% double track than a freight-only corridor at 19% double track.

Conclusions

Regression analysis is a powerful technique for comparing simulation results to determine, independent of the MLOS, changes in the capacity of different infrastructure configurations. Train delays decrease linearly with additional sections of double track when the traffic volume is constant. These trend lines can be predicted and used to develop response surface models in the typical form of exponential delay-volume relationships. As a hybrid track configuration transitions from single to double track under a constant MLOS, the incremental capacity gained from each added section of double track increases as more double track is added to the corridor. The simultaneous operation of freight and passenger trains on a heterogeneous line can reduce both overall capacity and the incremental capacity gained from each section of double track. The marginal loss in capacity from heterogeneous operation is greater on lines close to full double-track lines than hybrid track configurations that are closer to single-track lines. When large speed differentials are present between different types of trains, the speed differential may not be a significant delay-causing mechanism until most of the line is double track.

Acknowledgments

The authors thank Eric Wilson and others at Berkeley Simulation Software for the use of Rail Traffic Controller (RTC); Mark Dingler of CSX for technical insight; and Samantha Chadwick, Ivan Atanassov and Scott Schmidt of the University of Illinois at Urbana–Champaign (UIUC) Rail Transportation and Engineering Center (RailTEC) for their assistance in completing this research.

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: Partial support for this research was provided by the Association of American Railroads and the National University Rail (NURail) Center, a US DOT OST-R Tier-1 University Transportation Center. The first author was partially supported by a CN Research Fellowship in Railroad Engineering.

References

- Richards CW and Cobb T. Multiple-track main lines. *Trains* 2006; <http://trn.trains.com/railroads/railroad-maps/2010/01/free-preview-multiple-track-main-lines>.
- Lindfeldt O. From single to double track: effects of alternative extension measures. In: Brebbia CA, Tomii N, Mera JM, et al. (eds) *Computers in railways XIII: computer system design and operation in the railway and other transit systems*. Southampton, UK: WIT Press, 2012, p.313.
- Bing AJ, Beshers EW, Chzvez M and Simpson DP. *Guidebook for implementing passenger rail service on shared passenger and freight corridors*. Report 657, 2010. Ottawa, Canada: Canadian Transport Commission.
- Abril M, Barber F, Ingolotti L and Salido M. An assessment of railway capacity. *Transport Res, Part E: Logistics Transport Rev* 2008; 44(5): 774–806.
- Welch N and Gussow J. Expansion of Canadian National Railway's line capacity. *Interfaces* 1986; 16(1): 51–64.
- Kahn AM. *Railway capacity analysis and related methodology*. Report, 1979. Canadian Transport Commission, Ottawa, Canada: Canadian Transport Commission. <http://trid.trb.org/view.aspx?id=141733>.
- Vromans MJCM, Dekker R and Kroon LG. Reliability and heterogeneity of railway services. *Eur J Oper Res* 2006; 172: 647–665.
- Leilich RH. Application of simulation models in capacity constrained rail corridors. In: *The 30th winter simulation conference*, Washington, DC, 13–16 December 1998, pp.1125–1133. CA, USA: IEEE Computer Society Press Los Alamitos.
- Dingler M, Lai YC and Barkan CPL. Impact of train type heterogeneity on single-track railway capacity. *Transport Res Record: J Transport Res Board* 2009; 2117: 41–49.
- Sogin SL, Barkan CPL and Saat MR. Simulating the effects of higher speed passenger trains in single track freight networks. In: *The winter simulation conference*, Phoenix, AZ, 11–14 December 2011, pp.3679–3687. New York, NY: IEEE.
- Sogin SL, Barkan CPL, Lai YC and Saat MR. Impact of passenger trains in double track networks. In: *The joint rail conference*, Philadelphia, PA, 17–19 April 2012. New York, USA: ASME.
- Petersen E and Taylor A. Design of single-track rail line for high-speed trains. *Transport Res, Part A: General* 1987; 21(1): 47–57.
- Pawar SPS. *An analysis of single track high speed rail operation*. PhD Thesis, University of Birmingham, UK, 2011.
- Prokopy JC and Rubin RB. *Parametric analysis of railway line capacity*. Report DOT-FR-5014-2, 1975. Washington, DC: U.S. Department of Transportation.
- Krueger H. Parametric modeling in rail capacity planning. In: *The winter simulation conference*, Phoenix, AZ, December 1999, pp.1194–1200. New York, USA: ACM.
- Mitra S, Tolliver D and Mitra S. Estimation of railroad capacity using parametric methods. *J Transport Res Forum* 2010; 49(2): 111–126.
- Lai YC, Huang YA and Chu HY. Estimation of rail capacity using regression and neural network. *Neural Comput Appl* 2014; 25(7): 2067–2077.
- Mattsson LG. Railway capacity and train delay relationships. *Critical Infrastructure*. Springer Berlin Heidelberg, 2007, pp.129–150. <http://www.springer.com/us/book/9783540680550>.
- Wilson E. *Rail traffic controller (RTC)*. Berkeley, CA: Berkeley Simulation Software, 2012.
- Cambridge Systematics Inc. National rail freight infrastructure capacity and investment study, report, 2007.
- Galloway D. Train performance characteristics. Presented at the Transportation Research Board Workshop on Railroad Capacity and Corridor Planning, 2002.
- Federal Railroad Administration. *Rules, standards, and instructions governing the installation, inspection, maintenance, and repair of signal and train control systems, devices, and appliances*. Washington, DC: Federal Railroad Administration, 2000, pp.487–488.
- SAS Institute Inc. Using JMP 10. SAS Institute Inc. Cary, NC, 2012.