

Analyzing the Incremental Transition from Single to Double Track Railway Lines

Samuel L. Sogin*
University of Illinois
Department of Civil Engineering,
205 N Matthews, B118, Urbana, IL 61820
United States, e-mail: ssogin2@illinois.edu

Yung-Cheng Lai
National Taiwan University

C. Tyler Dick
University of Illinois

Christopher P.L. Barkan
University of Illinois

Abstract

Long term demand for rail freight transportation in North America is projected to increase considerably in the coming decades. Additionally, government agencies want to increase the speed and frequency of passenger trains operating on certain freight lines, further adding to demand for new capacity. A significant portion of the routes in the United States are single track with passing sidings. As traffic increases, additional trackage will be necessary to maintain network fluidity. Incremental additions to infrastructure should be placed to match traffic growth. The most likely approach will be addition of a second main track along portions of a route. This second track can be phased in over time creating hybrid track configurations. As the second track is installed, the operation will transition from single-track characteristics to double-track characteristics. Rail Traffic Controller (RTC) was used to simulate various hybrid track configuration under different operating conditions. In addition to the amount of second main track added, the analysis considered the interaction of traffic volume, traffic composition, and speed differential between train types. Adding sections of double track reduces train delay linearly under constant volume. Because of this linear relationship in delay, capacity will have a convex transition function from single to double track. These results will facilitate the development of an optimal incremental upgrade model for capacity expansion.

Keywords

Railways, simulation, shared corridors

1 Introduction

There are many factors that may determine how trains perform over a railway network. Railway simulation software continues to become more sophisticated in order to better emulate actual operations. The purpose of this study was to conduct a systematic analysis using controlled experiments focused on a subset of these factors.

Most of the United States railway network is single-track with passing sidings. Only 37-percent of mainlines with 10 million gross tons or more are multiple-mainline-track territory [1]. Future demand for increased freight and passenger rail service will require more capacity. Consequently, many of these single-track railway lines will need additional tracks to accommodate this demand. There are three basic approaches: adding

passing sidings, extending the siding length, and adding double track. Extending passing sidings enables longer freight trains and reduces passenger delays due to meets with other trains. Additional sidings would typically be installed in the longest single-track bottleneck sections. After these intermediate solutions are implemented, double track may be the most effective way to handle the additional traffic. This second mainline track can be phased in over time such that the amount installed is matched to the expected increases in rail traffic [2]. These intermediate phases with partial double tracking will be referred to in this paper as “hybrid” track configurations.

This analysis will focus on the capacity benefits of a single-track route transitioning 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 at different traffic levels. A typical North American single-track route was used as a baseline condition. Sections of double track were systemically added to the base condition by connecting pairs of pre-existing passing sidings. Train delay and capacity transition curves are then mathematically described. This procedure was used first with a homogeneous freight corridor. This was then compared to a shared corridor setting in which 25-percent of the total traffic was passenger trains. In order to differentiate between delay mechanisms, this shared corridor condition was simulated twice. The first run was with low-speed, high-priority passenger trains to determine the impact of priority. The second run 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 needed 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. The results presented in this paper are not intended to represent absolute predictive measurements for a particular set of conditions. Rather, they are meant to illustrate comparative effects under different conditions.

1.1 Background

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 the freight railway lines to provide passenger service. In most circumstances, the goal is that the level of service of the freight railway should be unaffected by the additional passenger traffic added to the route. This is usually accomplished by installing more track to mitigate potential delays. In some situations, the passenger agency may pay freight railroads for the slots that the passenger trains consume.

Experience and previous research have shown that double and single-track railway lines behave differently [4]. Single-track railways have considerably lower capacity than double-track lines. The primary reason for this reduction in capacity is due to trains traveling in opposite directions having to 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, theoretical capacity is primarily affected by the following distance between trains moving in the same direction. Double-track capacity is reduced if there are speed differentials, overtakes, or trains traveling against the current of traffic on the opposite direction track [5].

Previous studies of single track showed that adding a high-priority train to a freight network will increase average train delays more than simply adding another train that has similar characteristics to others operating on the line. Vromans et al used simulation

analyses to investigate strategies to improve passenger operations [6]. Leilich et al and Dingler et al used simulation analysis to evaluate the interactions between unit freight trains and high-priority container trains, and found a capacity loss due to the heterogeneous operations [7], [8]. Sogin et. al simulated single-track shared corridor conditions and determined that the maximum speed of the high-priority train had little relationship with train delay. Their analysis showed that delay distributions skewed to the right with none of the trains performing close to the minimum run time. Meets at sidings were cited as a primary delay mechanism [9]. Double-track configurations are more sensitive to speed differentials than on single track. A faster-high-priority train may need to use the second track in the opposing direction to overtake a slower train, thereby interfering with oncoming traffic on that track. Double-track configurations have delay distributions similar to exponential distributions with many trains operating close to the minimum run time [10].

There has been limited research on hybrid-track configurations. Petersen et al used simulation analysis to locate longer sidings in order to accommodate passenger trains on a freight line [11]. Additionally, Pawar et al used analytical models to determine the length of long-sidings in order to run a single-track, high speed railway without delays in meets [12]. Lindfeld compared partial double track to additional sidings and determined that it offers more flexibility to timetables and improves practical capacity more than additional sidings [2].

Regression modelling of train delays has been used in the past to quantify the effects of various operational factors on train delay. Prokopy and Rubin used single-track simulation results to develop a multivariate regression model [13]. Kruger developed a similar approach using an updated simulation model and also summarized his data with multivariate regression [14]. Both models were developed varying only one parameter at a time to isolate the effects of each. Mitra et al. developed an 8-variable regression model using simulation results for single-track lines, but their model did not consider interaction effects between variables [15]. Lai and Huang used regression and neural networks to model Rail Traffic Controller (RTC) simulation results from both a single and double-track network [16]. For both the single and double-track models, Lai and Huang used a full-factorial experimental design that analysed five factors at three different levels.

1.2 Delay as a Proxy for Capacity

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

Another approach to determining railway capacity is analysis of railway delay-volume curves. Using this method, train delay can be predicted as an exponential function of traffic volume, (Equation 1) [14,16,17]:

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

where

- D = Train delay
- V = Traffic volume
- A = Route constant
- k = Delay growth constant

A railway could then define the capacity of the line as the traffic volume (number of the trains per day) where the level of service (LOS) 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. For this paper, MLOS will be defined as the maximum average train delay tolerated by the railway operator, D_{max} . Under this definition for MLOS, Equation 1 can then be rearranged and solved for railway capacity.

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

Consider two different single-track routes that have different amounts of double-track sections installed. The delays from these two routes can be explained by Equation 1. Assuming that each route is operating at the traffic volume at the MLOS, D_{max} . By Equation 3a then the difference in capacities can be solved. The capacity difference of the capacities of these two lines is independent of the MLOS.

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

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

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

If the delay growth constants between the two routes are approximately equal then k will be approximately equal to 0. In this case, Equation 3c can be approximated by Equation 4. This may be a reasonable assumption when the types of traffic interactions between two infrastructure configurations are similar. However, a homogenous freight line may have a growth constant that is different with 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.

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

We presume that there is some functional relationship describing the relationship between capacity and the percentage of double track. Five hypothetical transition functions are shown in Figure 1. These five curves are all upward sloping assuming that capacity will always increase as more track is added. If capacity were measured on the y-

axis, then there would be an assumed positive relationship with the amount of double track installed. The shape of these curves may be different for different performance metrics. In this paper we will try and 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.

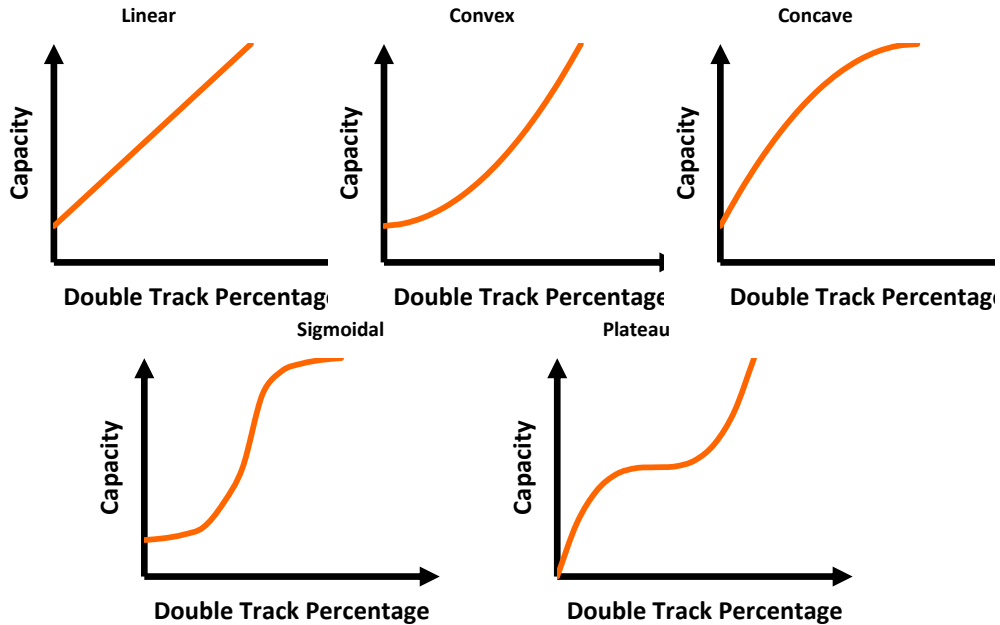


Figure 1: Potential shapes of transition functions from single to double track

2 Simulation Methodology

The hybrid track configuration experiment examined traffic volume, traffic mixture and the amount of 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 entry speeds, dense siding spacing, and capacity for long trains.

There are various strategies for how the second mainline track could be constructed in phases and distributed across a corridor. There will likely be sections of the route that cost more to construct than others. Based on a strategy of minimizing capital investment per kilometre of double track, the inexpensive sections of double track would be expected to be constructed before the expensive sections. However, if cost was not a factor in the decision making process, or if the cost of all double-track segments was relatively consistent, then there might be an optimal distribution of double track from an operating perspective. Two potential grouping strategies are illustrated in Figure 2: *alternate* and *split*.

Table 1: Route Parameter Guidelines

Parameter	Value
Length	386 kilometres
Bottleneck Length	12.9 kilometres
Siding Spacing	16.1 kilometres
Siding length	3,048 metres
Diverging turnout speed	72.4 km/h
Traffic control system	2-Block, 3-Aspect CTC
Average signal spacing	3.2 kilometres

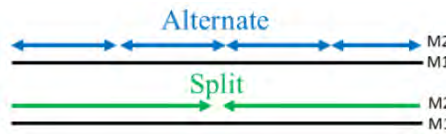


Figure 2: Double-track allocation Strategies

The *alternate* strategy is to pick four to six points on the line and build-out in both directions from these midpoints. This has the benefit of creating long sections of double track where two trains can meet without having either train stop at a siding. Additionally, these double-track sections may be long enough to achieve an overtake maneuver, where a faster train overtakes a slower train by using the opposite track. Another hybrid strategy is to *split* the double-track resource between the two terminals on the line and build towards the midpoint. The *split* strategy has the benefit of addressing potential bottleneck constraints at 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 being added to connect pairs of pre-existing sidings. When a siding becomes part of a section of mainline track, its track speed is upgraded to match the rest of the mainline. 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 sidings and second-mainline track. In this case, the baseline single-track configuration had 73.4-km of passing of passing sidings and therefore classified as 19-percent double track. This accounting of double-track percentage is the same as the ratio of double-track kilometres per route kilometre.

RTC, developed by Eric Wilson of Berkeley Simulation Software, is the *de facto* standard for railway simulation analyses in the United States [18]. Users include all the 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 taking into account allowable track-speeds, grade, curvature, and signalling systems. RTC will also modify train paths when trains are in conflict with each other, such as two trains simultaneously requesting use of the same section of track. Once a conflict is recognized, the logic reroutes and/or delays trains, as needed according to the priority of the trains. Trains are initially assigned user-defined priorities and departure times. As conflicts are resolved, train priorities are varied dynamically within user-defined bounds. 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 container trains over bulk commodity trains or passenger trains over freight trains. The architecture behind RTC is shown in Figure 3 [18].

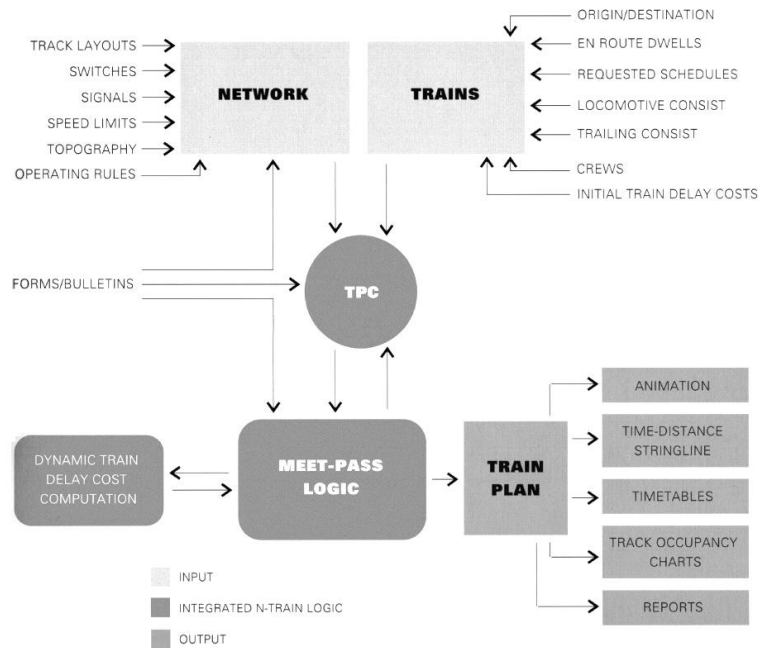


Figure 3: Architecture of Rail Traffic Controller [18]

An RTC simulation run for a particular scenario analyses five days of traffic in the corridor. Each simulation run is then repeated six times to yield performance statistics for 30 days of railway traffic for each of the runs in the experimental matrix. If a particular randomized run was infeasible for RTC to dispatch, then it is likely that one or more of the other six replicates was feasible and the scenario can still be used in the final analysis. Additionally, replication gave the opportunity for the dispatching algorithm to make different decisions with similar inputs to the model.

Although the infrastructure differs between cases to reflect 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 excess receiving and departure tracks. The double track is installed to minimize reconfiguring turnouts and control points of the signalling system. When a section of double-track connects a siding as mainline track, the turnout for that siding is reused as part of a future universal crossover leading into a future section of double track. The result is crossovers that are spaced at approximately 16-kilometre intervals on the double-track segments.

In addition to the physical and operating characteristics identified in Table 1, the schedule of traffic affects corridor performance. 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, and

other sources of variation and delay. Dispatchers are resolving conflicts between trains in real time instead of following a strict timetable. Different schedules of trains for the same infrastructure can show different average train delays. Consequently, assuming only one schedule for any given infrastructure may result in high experimental error, which will lead to bias in the results. In order to counteract error due to schedule bias, the departure time of each train is determined using a random uniform distribution over a 24 hour period. The randomized scheduling process is expected to create a range of schedules such that “good” low-delay schedules are averaged against “bad” high-delay schedules over the set of simulation replicates for each scenario. Stable averages can be achieved by averaging train performance over multiple days.

Each train in the simulation is based on the characteristics specified in Table 2. The freight train characteristics are based on the Cambridge Systematics National Rail Freight Infrastructure Capacity and Investment Study (2007) conducted for the Association of American Railroads (AAR) [19]. Freight car tonnages and lengths are based on averages for each car type. The power-to-ton ratios are based on experience and information from the Transportation Research Board Workshop on Railroad Capacity and Corridor Planning (2002) [20]. The passenger train is based on the 177-km/h operation between Chicago and Detroit. The acceleration curves from the simulation model were matched to GPS coordinate data from the Amtrak geometry car on this corridor. The passenger train stops were spaced at 52.1 kilometre intervals based on the current Amtrak station spacing on routes in California, Illinois, Washington, and Wisconsin. An 80-km/h maximum freight train speed is typical for a well maintained high-density mainline. Without in-cab signalling systems, the maximum speed of passenger trains is limited by regulation to 127-km/h in the U.S. [21]. Potential maximum speed upgrades on developing shared corridors across North America are 145-km/h and 177-km/h.

Table 2: Train Parameters for Simulation Model

	Freight Train	Passenger Train
Locomotives	x3 SD70	x2 P42
No. of Cars	115 hopper cars	8 Single level cars
Length (m)	1,928	152
Mass (kilotonne)	14.9	0.73
Power/Mass	0.78	15.4
		52.1 km between stops

Train delay is used as a proxy for capacity in this paper. While train delay is not a flow measurement, high train delays are indicative of a congested or saturated network. Train delays can be predicted by the traffic volume that can then give insight into the capacity of the railway line. Passenger traffic has a lower tolerance for delay compared to most freight traffic. These analyses will focus on freight train delays because these delays are most responsive to the identified factors. Passenger trains are shielded from experiencing very high delays due to their higher priority. Regression models of passenger train delays will be developed in the future.

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 [9,10]. The effect of priority was analysed by having a high priority passenger train travel at the same speed as the freight trains, 80-km/h. The effect of speed differential and priority acting together was represented by 177-km/h high priority passenger trains and 80-km/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 siding in a meet with two oncoming westbound freight trains. This can result in lower network delays than by splitting this conflict into two separate conflicts at two sidings. In single-track networks, passenger trains are often delayed by meets with other high priority passenger trains traveling in opposite directions because the siding length is only three kilometres.

3 Developing the Response Surface Model

The following analysis will focus on developing a response surface model based on the simulation data. The goal is to be able to predict the capacity of a line as a function of the amount of double track installed and the MLOS. The analysis in this section will show 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 will decrease linearly for each marginal section of double track installed 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 between 8 trains per day (TPD) and 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-percent) to complete-two-mainline track (100-percent). The linear reduction in train delay is greater with higher traffic levels than with lower traffic levels. Additionally, these trend lines are pivoting around approximately 100-percent double track.

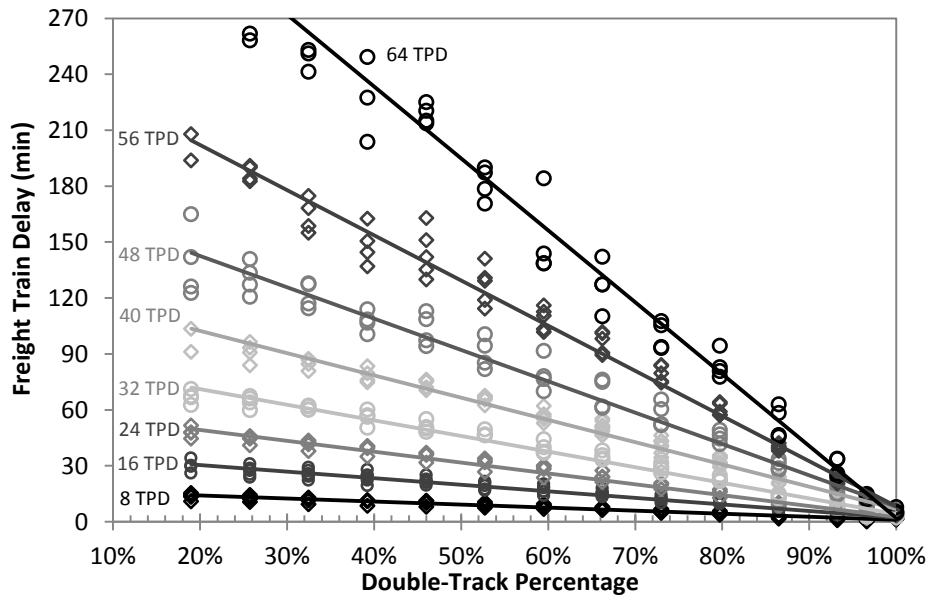


Figure 4: Train delays as a function of percentage of double at various traffic volumes

Each of the trend lines in Figure 4 can be described by slope and intercept parameters. If there are clear relationships between these parameters and the traffic volume on the line, then there can be a master equation that predicts delay 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 kilometre 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 centred on 19-percent, the single-track configuration. This point-slope format results in the intercept term relating 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 single track with zero 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-percent and indicates a small amount of triple-track.

$$D(V, x) = \gamma_0(V) - \gamma_1(V) \cdot (x - 19\%). \quad (5)$$

where

- $\gamma_0(V)$ = Single-track train delay as a function of traffic volume (intercept)
- $\gamma_1(V)$ = Reduction in train delay per double track-percentage point as a function of traffic volume (slope)
- x = 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 approximately full double track in Figure 4. Both the slope and single-track-delay 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 greater difficulty in deriving physical meaning from parameter estimates. Lastly, using an exponential relationship is more likely to simplify to an equation similar to Equation 1. If exponential relationships are assumed then Equation 5 becomes Equation 6.

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

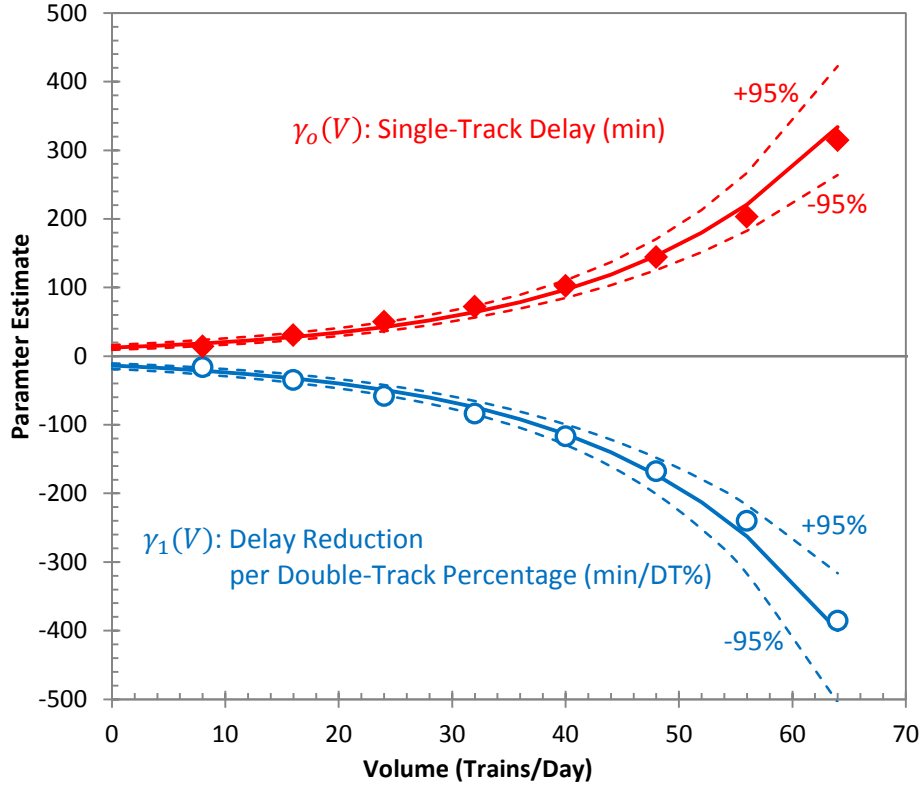


Figure 5: Linear parameter estimates of train delay reductions by adding sections of two-mainline-track. These parameter estimates are then predicted by exponential relationships

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

where

A_0 = Single track train delay constant

k_0 = Single track congestion factor

A_1 = Slope constant

k_2 = 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 7a and the slope by Equation 7b. These relationships were determined using a log-transformation and simple linear regression procedures in JMP [22]. The dashed lines represent 95-percent confidence bands around the mean response of the linear parameter estimate. Equations 7a and 7b can accurately predict the linear parameter estimates and be substituted into Equation 6 and yield Equation 8.

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

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

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

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 7a and 7b 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, non-linear regression can be used to arrive at the final four parameter estimates of Equation 6 without the loss in degrees of freedom. Not surprisingly, the non-linear regression platform yields more precise parameter estimates and results in a lower root mean square error (RMSE) of the original data.

Table 4: Comparison of the Hierarchical & Non-Linear Regression

		Statistic	Hierarchical	Non-Linear
		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

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-percent confidence intervals, there is clear overlap between these pairs of parameter estimates. While the model 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 centred on 19-percent 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 the line is then a function of the amount of double track installed and a delay standard, D_{max} .

$$D = (S_1 - S_2x)e^{kv} \quad (9)$$

where

S_1 = Single-track delay constant

S_2 = Delay mitigation constant

k = Congestion factor

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

Equation 10 is plotted in Figure 6 for different delay standards. The capacity improvement of the double-tracking is close to linear when the line is closer to a single track-configuration. As more double track is added and the line approaches full-two-main-track, the additional segments of second track yield increasingly greater capacity benefits. These capacity curves can help justify the last kilometre investments to complete the double-tracking of a line. These may be expensive tunnels, bridges, mountain passes or improvements in urban areas. The delay standard has more of an effect of 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).

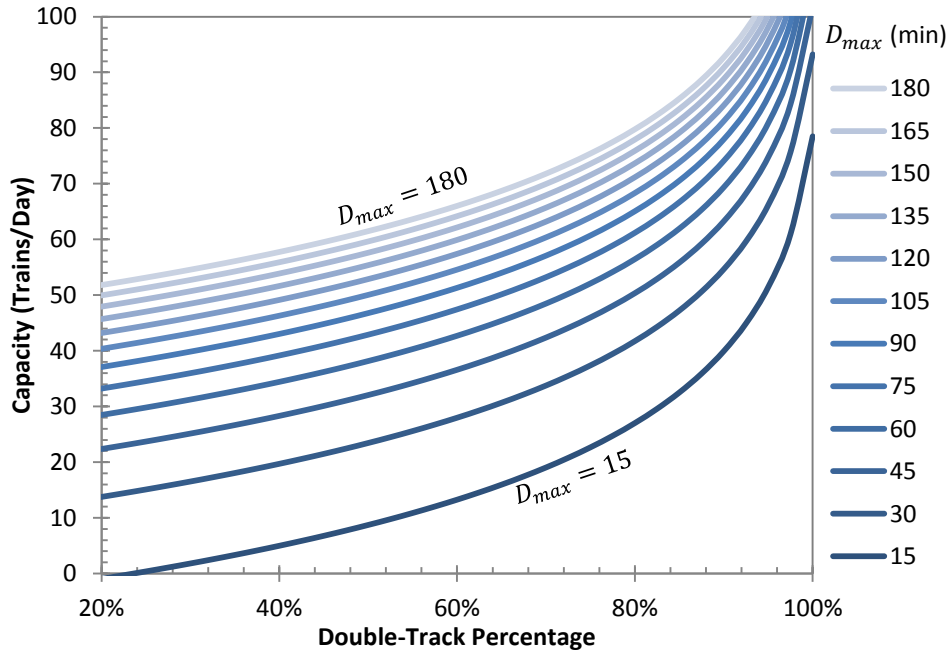


Figure 6: Capacity as a function of the amount of double track installed under different delay standards, D_{max} .

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 derivative of Equation 10 with respect to the double-track percentage which yields Equation 11. The implication of this 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-percent double track to 70-percent double track. Because Equation 11 does not include the delay standard, D_{max} , changing the delay standard will result in a constant change in capacity that is independent of the double-track percentage, x . For example, changing the delay standard from 30-minutes to 60-minutes will increase capacity by 14.2 TPD regardless of double-track percentage.

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

4 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 will be 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 trains per day will be considered in order to develop parameter estimates for Equation 9 for other conditions. The capacity of two or more configurations is compared by using Equation 10 at a MLOS set to 60-minute 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-percent double track and a 60-minute 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 will have 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 difference between these two allocation strategies for homogenous freight trains does 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 train per day (Figure 7). The parameters for these two scenarios are summarized in Figure 5. While there may not be much 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 configurations may show more benefit. If the model followed a strict timetable, then meets between trains can be planned to occur at sections of double track and take better advantage of the *alternate* configuration.

Table 5: Parameter Estimates for Different Allocation Strategies

	Alternate	Split	% Change
S_1	19.5206	18.4404	-5.53%
S_2	19.1490	18.0585	-5.69%
k	0.0471	0.0469	-0.42%

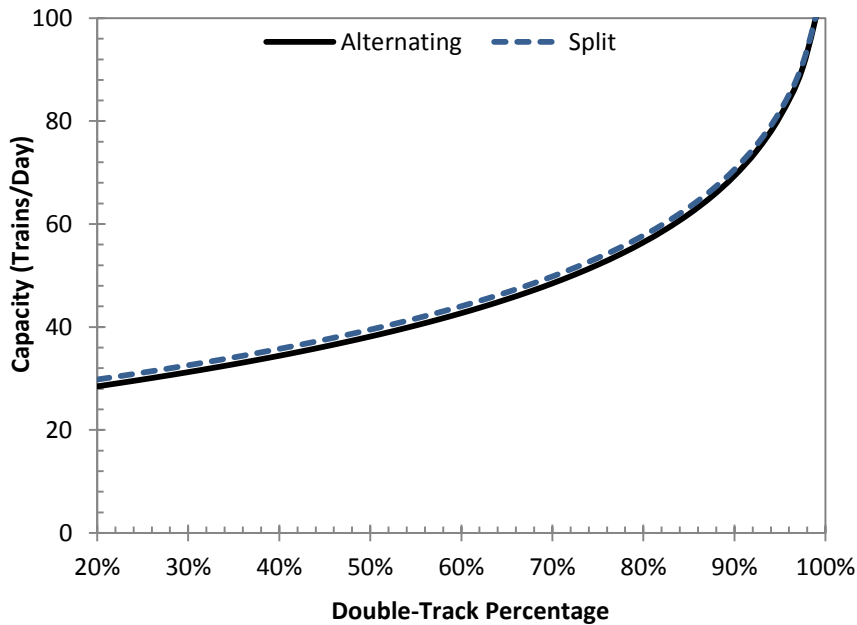


Figure 7: Capacity improvement under a 60 minute delay standard when sections of double track are allocated to the terminals (split) or at a collection of midpoints along the line (alternate).

We considered two potential mechanisms that might cause additional delay to freight trains from passenger trains on lines where they share trackage. The first is a priority differential between train types in which passenger trains are given preference in meet or pass conflicts. The second delay mechanism studied was speed differentials between train types. The effect of priority will be illustrated in a mixed traffic line where there are three freight trains per passenger train. Both trains will be limited to maximum speed of 80-km/h. The effect of speed and priority was evaluated by having 25-percent of the total traffic comprised of 177-km/h passenger trains. The double track was allocated using the *alternate* strategy.

A heterogeneous mixture of three freight trains per passenger train will by itself result in a capacity loss for any delay standard. Additionally, it will take more double-track to mitigate traffic increases. Having priority trains only manifests a change relative to the base case by having a higher k coefficient, indicating higher sensitivity to traffic increases. The higher speed passenger trains have a k coefficient on par with the freight-only case but also have higher S_1 and S_2 coefficients. The capacity curves for 60-minute MLOS are plotted in Figure 8. The difference between 80-km/h and 177-km/h is very small between 20-percent and 80-percent double track. In the 80 to 100-percent double-track range, there is a divergence between the two passenger train interference curves, where speed differentials start to reduce capacity.

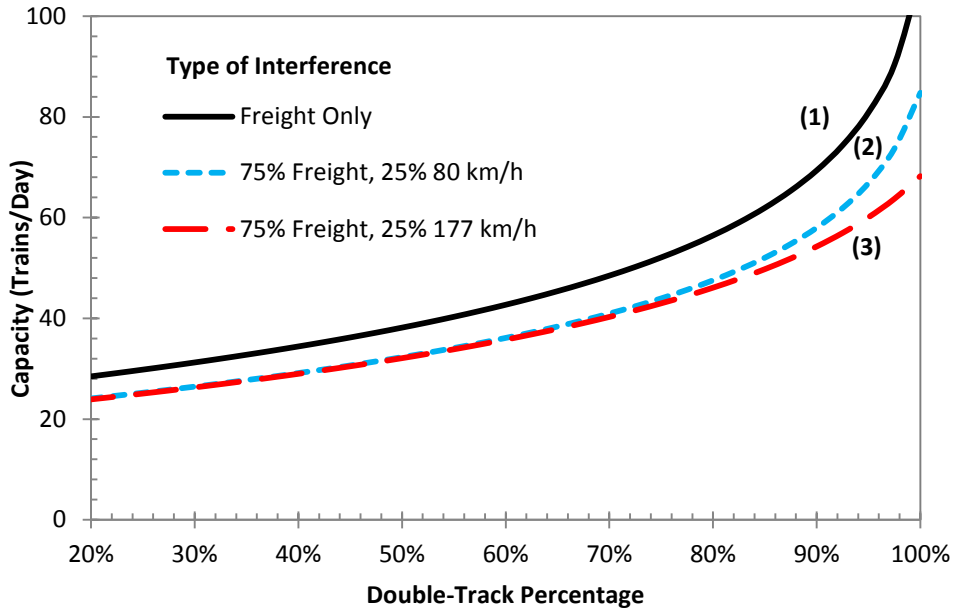


Figure 8: Change in capacity by installing sections of double track with a 60-minute delay standard. Where (1) the traffic is 100-percent freight trains, (2) the traffic is 75-percent 80-km/h freight trains and 25-percent 80-km/h passenger trains, and (3) the traffic 75-percent 80-km/h freight trains and 25-percent 177-km/h passenger trains.

Table 6: Parameter Estimates of Equation (9) for Different Traffic Types

	Freight Only	80 km/h Passenger Trains	177 km/h Passenger Trains
S_1	19.5206	19.9317	22.4534
S_2	19.149	19.3509	20.4052
k	0.0471	0.0547	0.0495

The loss in capacity due to 177-km/h passenger trains in a freight corridor under an average 60-minute MLOS is illustrated in Figure 9. The capacity loss due 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 80-km/h passenger trains, to the total loss from 177-km/h passenger train interference. The effect of priority differential accounts for 96-percent of the capacity loss between 19-and-55-percent double track. At full 100-percent double track, the speed differential mechanism accounts for 41.7-percent of the loss in capacity, and priority accounts for 58.3-percent.

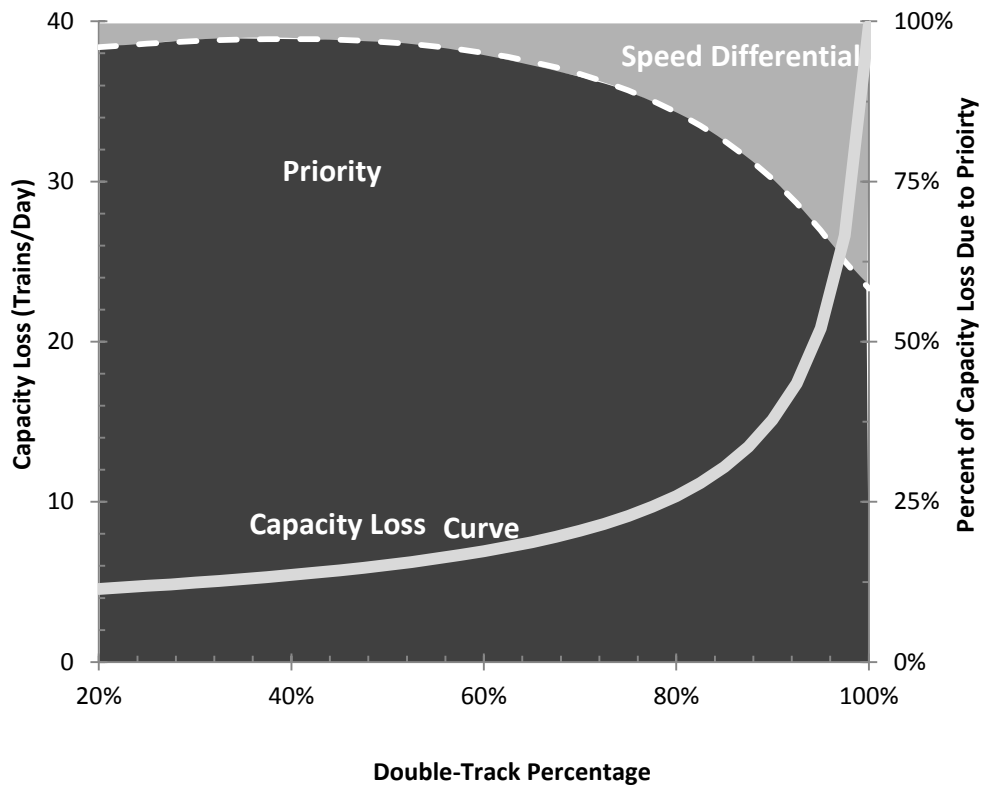


Figure 9: Capacity loss of heterogeneous operations over homogenous operations highlighted by relative effect of delay mechanics

The freight train capacity loss on a higher-speed passenger, mixed use corridor is even greater than that shown in Figure 9 because it assumes 25-percent 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-percent freight and 25-percent 177-km/h passenger trains. The initial capital investment mitigates the additional delays to the original freight trains and improves capacity by 33 percent to accommodate the additional passenger trains. If the freight line was originally single track, then this initial investment will be to upgrade the line from 19-percent double track to 65-percent double track under a 60-minute MLOS. The amount of double track needed to accommodate the passenger trains was calculated using Equation 13. This substantial investment in capacity does not benefit the freight railway; it simply allows them to maintain their current level of service. In this 75-percent-freight, mixed-use corridor, any future growth in both freight and passenger volume will require additional capacity investment with at 65-percent 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 train types were eliminated then the initial capacity investment would

be to upgrade the corridor to 50-percent double track. With 80-km/h passenger trains, there is also more freight capacity available in a full two-track build out than with 177-km/h passenger trains.

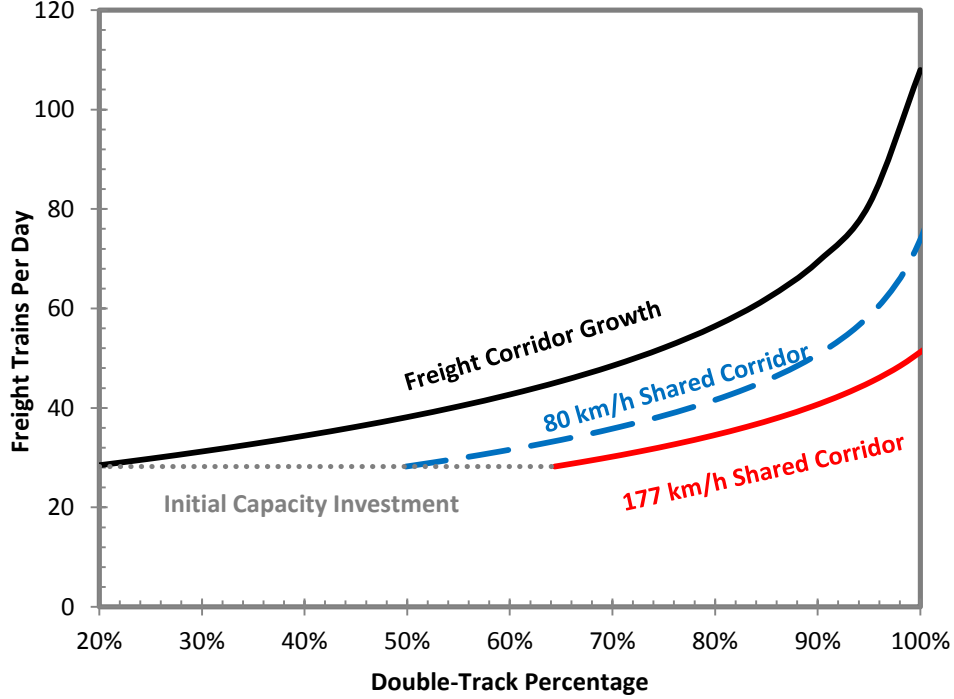


Figure 10: Freight train capacity grown under a 60 minute MLOS for a freight corridor and a shared corridor comprised of 75% freight and 25% passenger trains.

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, then the next time freight demand increases, the freight railway will receive a higher return on capacity per track-kilometre installed. This benefit occurs because the freight railway is now on the 177-km/h shared corridor curve in Figure 10, which is steeper at 65-percent double track than a freight only corridor at 19-percent double track.

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

where

x_m = Level of double tracking to mitigate addition of passenger service

S_{1p} = Shared corridor single-track delay constant

S_{2p} = Shared corridor delay mitigation constant

k_p = Shared corridor congestion factor

V_o = Initial freight corridor volume (capacity at D_{max})

φ = Freight train percentage of the total traffic of the planned shared corridor

5 Future Work

Equation 9 is a powerful model for relating train delays to traffic volume and the percentages of double-track. Further manipulation of various operating conditions can provide further insight on how the parameter estimates of Equation 9 change. For example, the three parameter estimates could be related to various levels of passenger train speeds. The derivation of Equation 9 depends heavily on the linear relationship between double-track percentage and train delay under constant volume. In this analysis, the sidings were all evenly spaced at 16.1-km. If there was a random distribution of distances between sidings than these trend lines may not be linear. Further investigation into this assumption is needed, as the capacity benefits of connecting longer bottleneck sections may be disproportionately greater than connecting shorter sections. Additionally, there are other strategies on how to allocate sections of double track across a corridor beyond the ones identified in Figure 2. RTC has limited ability to model terminal-mainline interactions. In this analysis, the terminal was designed large enough to minimize these effects; however if larger terminal-mainline interactions were designed into the model then these effects may dictate a greater effect of the allocation strategy.

6 Conclusion

Regression analysis is a powerful technique for comparing simulation results to determine changes in capacities of different infrastructure configurations that can yield results independent of the MLOS. Train delays decrease linearly with additional sections of double track when the volume is constant. These trend-lines can be predicted and can be 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 section of double track added increases as more double track is added to the corridor. The simultaneous operation of freight and passenger trains on a heterogeneous line can reduce 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 than hybrid track configurations that are closer to single-track lines. When large speed differentials are present between train types, the speed differential may not be a significant delay causing mechanism until most of the line is double tracked.

Acknowledgements

Partial support for this research was provided by grants from the Association of American Railroads and the National University Rail (NURail) Center, a US DOT RITA University Transportation Center. The corresponding author was partially supported by the CN Research Fellowship in Railroad Engineering. The authors thank Eric Wilson and others of Berkeley Simulation Software for the use of Rail Traffic Controller (RTC); Mark Dinger 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.

References

- [1] C. W. Richards and T. Cobb, "Multiple-Track Main Lines," *Trains*, Jan-2006.
- [2] O. Lindfeldt, "From Single to Double Track: Effects of Alternative Extension Measures," *Computers in Railways XIII: Computer System Design and Operation in the Railway and Other Transit Systems*, p. 313, 2012.
- [3] A. J. Bing, E. W. Beshers, M. Chzvez, and D. P. Simpson, "Guidebook for Implementing Passenger Rail Service on Shared Passenger and Freight Corridors," Washington D.C., 2010.
- [4] M. Abril, F. Barber, L. Ingolotti, and M. Salido, "An Assessment of Railway Capacity," *Transportation Research Part E: Logistics and Transportation Review*, vol. 44, no. 5, pp. 774–806, 2008.
- [5] A. M. Kahn, "Railway Capacity Analysis and Related Methodology," Ottawa, 1979.
- [6] M. J. C. M. Vromans, R. Dekker, and L. G. Kroon, "Reliability and heterogeneity of railway services," *European Journal Of Operational Research*, vol. 172, pp. 647–665, 2006.
- [7] R. H. Leilich, "Application of Simulation Models in Capacity Constrained Rail Corridors," in *Proceedings of the 30th conference on Winter simulation*, 1998, pp. 1125–1133.
- [8] M. Dingler, Y.-C. Lai, and C. P. L. Barkan, "Impact of Train Type Heterogeneity on Single-Track Railway Capacity," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 640, no. 2117, pp. 41–49, 2009.
- [9] S. L. Sogin, C. P. L. Barkan, and M. R. Saat, "Simulating the Effects of Higher Speed Passenger Trains in Single Track Freight Networks," in *Proceedings of the 2011 Winter Simulation Conference*, 2011, pp. 3679–3687.
- [10] S. L. Sogin, C. P. L. Barkan, Y.-C. Lai, and M. R. Saat, "Impact of Passenger Trains in Double Track Networks," in *Proceedings of the 2012 Joint Rail Conference*, 2012.
- [11] E. Petersen and A. Taylor, "Design of Single-Track Rail Line for High-Speed Trains," *Transportation Research Part A: General*, vol. 21, no. 1, 1987.
- [12] S. P. S. Pawar, "An Analysis of Single Track High Speed Rail Operation," University of Birmingham, 2011.
- [13] J. C. Prokopy and R. B. Rubin, "Parametric Analysis of Railway Line Capacity," Federal Railway Administration, Washington, D.C., 1975.
- [14] H. Krueger, "Parametric Modeling in Rail Capacity Planning," in *Proceedings of the 1999 Winter Simulation Conference*, 1999, pp. 1194–1200.
- [15] S. Mitra, D. Tolliver, and S. Mitra, "Estimation of Railroad Capacity Using Parametric Methods," vol. 49, no. 2, 2010.
- [16] Y.-C. Lai and Y.-A. Huang, "Estimation of Single and Double-Track Capacity with Parametric Models," *Transportation Research Board: 91st Annual Meeting*, pp. 1–22, 2012.
- [17] L.-G. Mattsson, "Railway Capacity and Train Delay Relationships," *Critical Infrastructure*, 2007.
- [18] E. Willson, "Rail Traffic Controller (RTC)." Berkeley Simulation Software, 2012.
- [19] Cambridge Systematics, "National Rail Freight Infrastructure Capacity and Investment Study," 2007.
- [20] D. Galloway, "Train Performance Characteristics," in *Transportation Research Board Workshop on Railroad Capacity and Corridor Planning*, 2002.
- [21] Federal Railroad Administration, *Rules, standards, and instructions governing the installation, inspection, maintenance, and repair of signal and train control systems, devices, and appliances*. 2000, pp. 487–488.
- [22] "JMP." SAS Institute INC., Cary, NC.