Relative train length and the infrastructure required to mitigate delays from operating combinations of normal and over-length freight trains on single-track railway lines in North America

C Tyler Dick [ID], Ivan Atanassov, F Bradford Kippen III
and Darkhan Mussanov

Abstract

Distributed power locomotives have facilitated longer heavy-haul freight trains that improve the efficiency of railway operations. In North America, where the majority of mainlines are single track, the potential operational and economic advantages of long trains are limited by the inadequate length of many existing passing sidings (passing loops). To alleviate the challenge of operating trains that exceed the length of passing sidings, railways preserve the mainline capacity by extending passing sidings. However, industry practitioners rarely optimize the extent of infrastructure investment for the volume of over-length train traffic on a particular route. This paper investigates how different combinations of normal and over-length trains, and their relative lengths, relate to the number of siding extensions necessary to mitigate the delay performance of over-length train operation on a single-track rail corridor. The experiments used Rail Traffic Controller simulation software to determine train delay for various combinations of short and long train lengths under different directional distributions of a given daily railcar throughput volume. Simulation results suggest a relationship between the ratio of train lengths and the infrastructure expansion required to eliminate the delay introduced by operating over-length trains on the initial route. Over-length trains exhibit delay benefits from siding extensions while short trains are relatively insensitive to the expanded infrastructure. Assigning directional preference to over-length trains improves the overall average long-train delay at the expense of delay to short trains. These results will allow railway practitioners to make more informed decisions on the optimal incremental capital expansion strategy for the operation of over-length trains.

Introduction

Longer freight trains achieve economies of scale that improve fuel consumption, reduce operating crew costs, and increase mainline capacity by using fewer trains to move the same volume of freight.[1–3] Outside of closed-loop mine-to-port operations, specialized shipments of single commodities in bulk "unit trains" of approximately 100 railcars began to evolve in the United States during the 1960s.[4] Various regulations on crew sizes and rates for multiple-car shipments limited many economic benefits of longer trains prior to deregulation of U.S. railroads in 1980. Practical train lengths were also limited by the

mechanical strength of couplers and locomotive adhesion. The advent of distributed power in the 1980s and alternating current traction locomotives in the 1990s allowed bulk unit train lengths beyond 120 railcars.[5] Thus, longer freight trains of over 100 railcars are a relatively new phenomenon on many parts of the

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

Corresponding author:
C Tyler Dick, Department of Civil and Environmental Engineering,
University of Illinois at Urbana-Champaign, Champaign, USA.
Email: ctdick@illinois.edu

North American rail network. In 1980, the average length of freight trains in the western United States was 68.9 railcars. Twenty years later in 2000, this had only increased to 72.5 railcars. In 2010, the average train length grew to 81.5 railcars as railways began to operate trains of 150 railcars on certain corridors.[6] Longer trains are not limited to bulk commodities as railroads have implemented 14,000 ft (4300 m) intermodal trains on transcontinental corridors.[7]

Widespread implementation of longer trains in North America is limited by existing track infrastructure layouts.[8] Approximately 70% of principal North American mainlines are single track with passing sidings (also known as loops or passing loops) where two tracks exist for only a short distance.[9] Unlike double track, in order for two trains traveling in opposite directions to meet and pass each other in an efficient manner at a passing siding on single track, at least one of the trains must be shorter than the length of the passing siding.[10] The length of typical existing passing sidings on single track in North America ranges from 6000 to 7500 ft (1830 to 2285 m) to accommodate trains of 100–120 railcars.[11] Most newly constructed passing siding projects are in the range of 9000–10,000 ft (2745–3050 m) to support the operation of trains with 150 railcars and seven distributed power locomotives.

Trains that exceed the length of passing sidings are referred to as "over-length trains" and their operations are restricted to avoid a meet between two over-length trains at a siding sized for shorter trains. To accommodate meets between over-length trains, railways invest capital to extend the length of passing sidings.[12] However, industry practitioners rarely optimize the number of siding extensions for the number of over-length trains on a particular route as they are reluctant to operate multiple over-length trains until nearly all passing sidings are extended. Thus, the 100-car length of many existing passing sidings effectively sets an upper bound for train lengths on the general North American railway network.

Contemporary utilization of 150-car unit trains and its relationship to the length of passing sidings has been addressed by both academia and industry. Grimes advised that long haul grain and coal movements should utilize maximum train sizes to minimize operating costs but that siding length and railroad policy can limit train size or the particular routes used by longer trains.[13] The operational and economic benefits of longer unit trains on one Class 1 railway were documented by Newman et al.[14] Barton and McWha[15] discussed the operational gains of longer freight trains and described the need for extended passing sidings to support utilization of longer 12,000 ft freight trains by several North American Class I railroads. Martland[16] noted the inability of existing passing sidings to support the operation of long trains and estimated that two-thirds of unit trains in operation are "length-limited"

by passing sidings. The relationship between siding length and the maximum practical train length on a mainline segment is demonstrated by Dick and Clayton; at the time of writing, the typical passing siding in Western Canada on Canadian Pacific Railway (CP) was several thousand feet longer than those on competitor Canadian National Railway (CN) and trains on CP were correspondingly longer than those on CN. To counter its siding-length disadvantage while avoiding the operational challenge of meets between over-length trains, CN began to run trains of 150 railcars (9000 ft or 2745 m, in length) in a single direction.[17]

Previous work conducted by the authors concluded that the introduction of over-length trains on a route without any extended passing sidings greatly increased average train delay.[18,19] For the combination of "short" 100-car trains and "long" 150-car trains, only one half of the passing sidings on a representative mainline need to be lengthened to mitigate the delay increase from the over-length trains.

The research presented in this paper aims to further advance the understanding of over-length train operation by generalizing the previous work over a more comprehensive set of short and long train lengths. Specifically, this paper investigates how the relative lengths of short and long trains, as quantified by the "train length ratio," relates to the number of siding extensions required to mitigate the delay from over-length trains on a single-track rail corridor. In addition, this research seeks to understand how the delay performance of long and short trains differs and changes as infrastructure is expanded and passing sidings are extended on a single-track rail corridor.

To meet these objectives, the research team conducted simulation software experiments and analyzed resulting train delays for various combinations of short and long train lengths under different directional distributions of a given daily railcar throughput volume. The results of the experiments increase knowledge of the fundamental relationships between the relative lengths of long and short trains, passing siding length, and train delay. Improved understanding of this relationship can lead to more informed decisions regarding efforts to extend existing passing sidings (or construct additional longer sidings) required to implement over-length train operations. Railway practitioners can make more efficient use of limited capital resources by better matching the number of long sidings to the relative size and quantity of over-length trains on a particular route.

Methodology

The technical approach of this study extends the examination of over-length trains previously conducted by the authors.[18,19] While the underlying simulation framework parallels the previous work, methodological advances are made by consideration

of multiple combinations of train lengths to vary the "train length ratio," and through consideration of train-type delays in addition to average train delay.

Train length and replacement ratio

This paper frequently refers to the concept of a "train length ratio" and "train replacement ratio." This research defines the train length ratio as the ratio between the length of "long" over-length trains on a mainline and the length of normal "short" trains on the same route (equation (1))

$$Train\ Length\ Ratio = \frac{Long\ Train\ Length}{Short\ Train\ Length} \quad (1)$$

If a mainline normally operates short trains of 100 railcars but is transitioning to long trains of 150 railcars, two long trains can move an amount of freight (railcars) equal to three short trains, yielding a 3:2 train replacement ratio. The interpretation of the train replacement ratio is that, for a fixed traffic volume, three short trains can be replaced by two long trains to reduce the total train count. A larger disparity between long and short train sizes results in a larger train length ratio and train replacement ratio. Train length ratios closer to 1 correspond to small increases in train length over normal operations.

Train delay

North American railroads monitor the performance and capacity of mainlines through various metrics, with train delay being the most common.[20] Consistent with this industry practice, this paper uses average train delay to evaluate and compare the relative performance of different combinations of short and long trains on a representative single-track railway corridor.

Calculation of train delay for North American freight operations is complicated by the use of improvised operations where trains do not have a detailed preplanned timetable of intermediate stops.[16] Without a detailed timetable, train conflicts must be resolved by train dispatchers in real time, creating schedule flexibility and additional dwell time on the route.[21] Under flexible freight operations, train delay is defined as the difference between a particular train's minimum running time and its actual running time (equation (2))

$$Train\ Delay = Actual\ Running\ Time \\ - Minimum\ Running\ Time \quad (2)$$

The minimum running time is the time required for a train to traverse a mainline segment with no stops due to meets with conflicting trains, while obeying all maximum authorized speeds and allowing for train acceleration and braking performance. The actual running time is the actual time elapsed while a train traverses the same mainline segment in the presence of train traffic, including the time the train is in motion and the time the train is stopped waiting for other trains or unexpected sources of delay. By this definition, train delay includes the time a train is stopped waiting for other trains. This definition is different from other international definitions of train delay where scheduled time spent dwelling in stations for planned trains meets is not considered as delay.[22] Under international definitions of train performance for scheduled timetable operations, train delay is analogous to "train excess time."

Rail Traffic Controller (RTC)

The train delay response for each experiment scenario described in subsequent sections is determined via RTC software. RTC is widely used by the North American rail industry, including, but not limited to, Amtrak, Class I railroads, and consultants.[23] RTC simulates dispatcher decisions in guiding trains along the specific routes to resolve meet and pass conflicts. General RTC model inputs include details of the track layout, curvature, grades, train characteristics, and signal control system.

To determine the corresponding train delay response, five days of train operations are simulated in RTC for each unique combination of number of siding extensions, train length ratio, and percentage of over-length trains in the experiment design. To consider variability in train departure times according to flexible North American operations, 25 days of train delay data are generated by replicating each five-day simulation five times. Each replicate represents a distinct freight train operating pattern where each train departs its respective terminal randomly within a 24 h time period. Long trains and short trains are distributed randomly within this pattern according to a uniform probability distribution; no efforts are made to fleet the trains by length. The randomized train departures allow the simulations to capture different patterns and sequences of train lengths and their associated "knock-on" or cascading train delay impacts.

The delay accumulated by individual trains over the 25 days of data is normalized by the number of accumulated train-miles (train-kilometers) to develop a train delay response for that scenario within the experiment design. The average train delay for all trains in a scenario, or long and short trains totaled separately, is plotted as a single data point in the results. Although commonly used in North America, a shortcoming of using average train delay as the key performance metric is that it does not describe the overall distribution of train delay or the delay experienced by the worst-performing trains that may drive railroad business decisions. Related research conducted by the coauthors and colleagues have examined the overall distribution of train delay under

various North American single-track operating conditions.[24–26] Differences in average train delay by train type will be considered in "Relative delay to short and long trains" section.

Route and train characteristics

The same baseline route infrastructure was used for all scenarios with route parameters selected to be representative of North American freight infrastructure and operating conditions. The route consists of 240 mile (386 km) of single-track mainline with terminals at each end. To eliminate irregular spacing of passing sidings as a source of variation,[27] passing sidings are located at a regular center-to-center spacing of 10 mile (16 km) (Table 1). Two types of sidings are included: shorter-length passing sidings to represent current conditions and a second longer-length passing siding to represent a siding that has been extended through capital investment. In a given scenario, the baseline route infrastructure is altered by extending a certain number (or percent) of short passing sidings to the longer length according to the experiment design.

Traffic volume is fixed at 2400 railcars per day with 50% of the railcars moving in each direction. All trains on the route are freight trains with equal priority and maximum authorized speed. Trains are either "short" or "long" relative to the initial passing siding length of 1.25 mile (2.01 km). Different short and long

train lengths are used in combination to achieve the train length ratios specified in the experiment design. The number of 4300 hp (3207 kW) diesel-electric locomotives assigned to a particular train is proportional to its length, with two locomotives on short trains and three on long trains. Adding power to the longer trains helps provide consistent acceleration performance between train sizes. Inconsistent train performance may have resulted in additional congestion and delay due to slow acceleration of longer trains, potentially confounding the simulation results.

The ability of heavy-haul unit train inspection facilities, intermediate staging yards, and loading and unloading terminals to support the operation of longer trains is not considered in this paper. Although this paper focuses on mainline operations of over-length trains on single track, additional yard and terminal investments may be required to extend the length of turning loops and yard tracks to support longer trains.

Experiment design

Three variable factors are included in the RTC simulation experiment design: percent long sidings, percent long trains, and train length ratio (Table 2).

"Percent long sidings" is the percent of the mainline passing sidings that are extended from the base length of 1.25 mile (2.01 km), designed for a maximum of 100 railcars, to a length of 2.00 mile (3.2 km) to exceed the 120- or 150-car length of the long trains. The experiment design includes various levels of incremental expansion from the base case with no long sidings to the case where all sidings are extended to the longer length. An idealized strategy was adopted when selecting which sidings along the route to extend. Passing siding extensions were always distributed evenly along the length of the route so that the route remained balanced from the perspective of passing siding length.

"Percent long trains" is the percent of the daily railcar volume on the mainline transported in long trains. It is also the percent of short trains that have been replaced by long trains to move the same number of railcars (volume of freight). For example, consider the combination of 150-car long trains and 100-car

Table 1. Simulated route and train characteristics.

Parameter	Value
Length of route	240 mile (386 km)
Siding spacing	10 mile (16 km)
Number of sidings	23
Initial "short" siding length	1.25 mile (2.01 km)
Extended "long" siding length	2 mile (3.22 km)
Traffic volume	2400 railcars per day
Traffic composition	100% freight
Maximum authorized speed	50 mph (80.5 kph)
Turnout diverging route speed	45 mph (72.4 kph)
Operating protocol	2-block, 3-aspect Centralized Traffic Control,
Locomotives	4300 hp (3207 kW) diesel-electric (two per short train or three per long train)
Short train consist	50, 75, or 100 railcars
Short train length	2750, 4125, 5500 ft (838, 1257, 1676 m)
Long train consist	120 or 150 railcars
Long train length	6600 or 8250 ft (2012, 2515 m)

Table 2. Simulation experiment design factors and levels.

Variable	Values
Percent long sidings	0, 4, 9, 13, 22, 30, 48, 52, 70, 78, 87, 91, 96, 100
Percent long trains	0, 25, 50, 75
Train length ratio	6:5 (120-car long and 100-car short)
	3:2 (150-car long and 100-car short)
	2:1 (150-car long and 75-car short)
	3:1 (150-car long and 50-car short)

short trains. The traffic level of 2400 railcars per day requires 24 short trains per day to move the given freight volume. For 50% long trains, 1200 railcars move in long trains and 1200 railcars move in short trains. This is equivalent to eight long 150-car trains and 12 short 100-car trains. The percent long trains factor is the number of railcars being moved in long trains divided by the total railcar throughput, rather than just the ratio of long trains to total number of trains. The experiment design includes four levels of percent long trains: 0 (all short trains), 25, 50, and 75% long trains. Where long trains are present in the main experiment design, the long trains operate bidirectionally, with half of the long trains operating in each direction to maintain directional balance. For example, in the scenario described above with a total of eight long trains and 12 short trains, four long trains and six short trains operate in each direction. An even directional distribution is consistent with heavy-haul unit train operations where complete load–empty cycles are made with a fixed train consist, limiting the ability of a railway to only operate long trains in a single direction.

To consider different train length ratios, the experiment design uses various combinations of trains with 50, 75, 100, 120, and 150 railcars. The cases of 6:5 (100-car short and 120-car long trains) and 3:2 (100-car short and 150-car long trains) length ratios are representative of typical North American conditions as heavy-haul operators increase the length of unit trains. To examine a greater range of train length ratios without simulating extremely long trains of 300 railcars, the cases with 2:1 and 3:1 length ratios consider short train lengths with fewer railcars than typically operated under current conditions.

To provide data for analysis, RTC simulations were completed for all 224 factorial combinations of the three variables at their various factor levels and a traffic volume of 2400 railcars per day.

To investigate the differences in train delay between short and long trains as increasing numbers of sidings are extended along the route, an additional set of simulations were conducted at a traffic volume of 3600 railcars per day, 50% long trains, and 3:2 length ratio (150-car long trains and 100-car short trains). To contrast against the bidirectional scenarios where the same number of long trains operates in each direction, an additional set of "unidirectional" simulations were conducted with all long trains operating in the same direction. The goal of these additional simulations was to investigate if the particular efficiencies afforded by directional operation of over-length trains identified by the author's earlier work[18,19] (namely elimination of meets between two long trains) alter the distribution of train delay between short and long trains.

Results

Train length ratio and required number of long sidings

The RTC simulation results were analyzed to calculate the average train delay for each experiment scenario. Average train delay was examined for different combinations of infrastructure investment (percent long sidings) and degree of long train operations (percent long trains) at each train length ratio (Figures 1 to 4). The condition with zero percent long trains (i.e. all traffic moving in short trains) is largely insensitive to the number of extended passing sidings along the route. This is expected as short train meets can take place in any passing sidings and there is no benefit from extended passing sidings other than slightly decreasing the distance between sidings. Thus, the zero percent long trains' trend line serves as a baseline average train delay for each train length ratio. The relative train delay performance of operations with different percent long trains at a given percent long sidings can be evaluated relative to this baseline. The delay-infrastructure curves for the largest train length ratio (Figure 3) exhibit greater fluctuations and
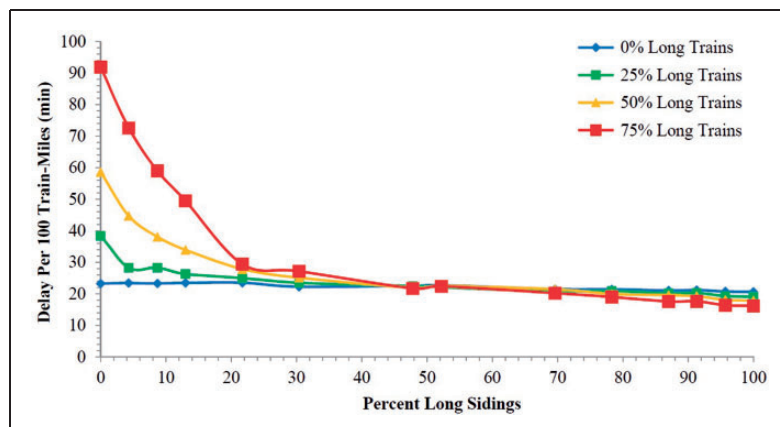

Figure 1. Average train delay for combinations of percent long trains and percent long sidings with 3:2 length ratio.
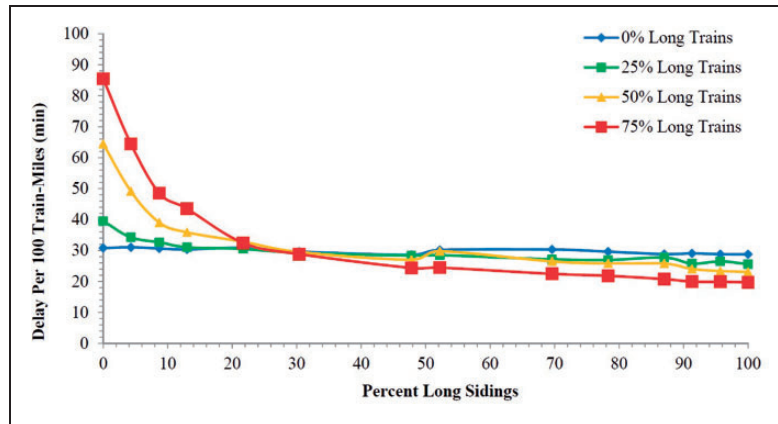
Figure 2. Average train delay for combinations of percent long trains and percent long sidings with 2:1 length ratio.
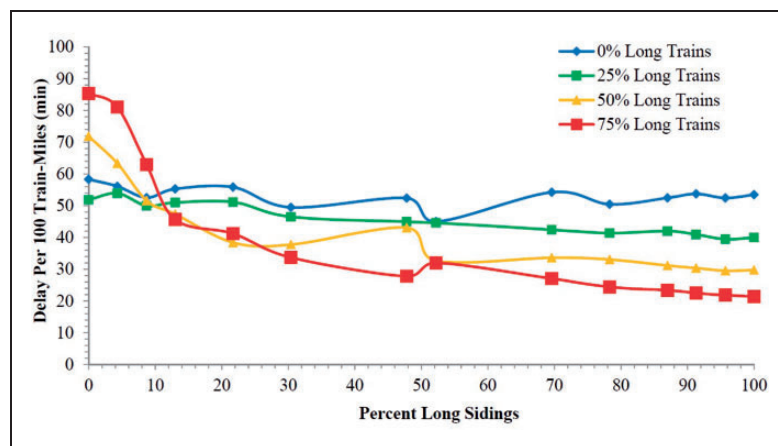

Figure 3. Average train delay for combinations of percent long trains and percent long sidings with 3:1 length ratio.
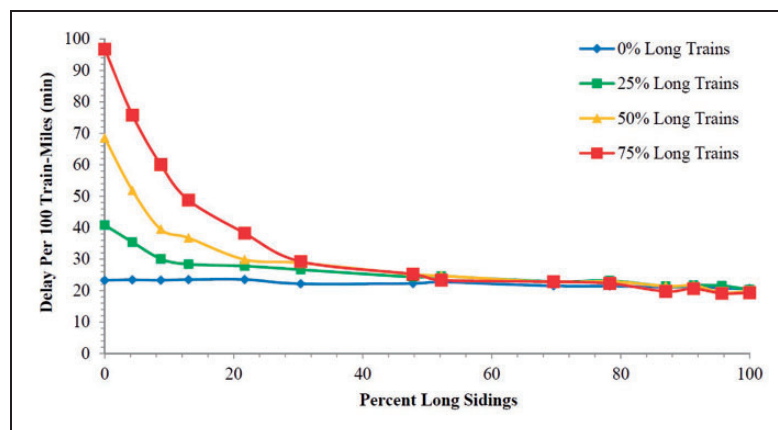

Figure 4. Average train delay for combinations of percent long trains and percent long sidings with 6:5 length ratio.

a nonlinear baseline because there are a greater number of trains on the route and certain random sequences of train departures can produce large cascading "knock-on" train delays.

Given the position of the baseline delay for zero percent long trains, train delay for a 3:2 length ratio (150-car long trains and 100-car short trains) fall into

three natural "zones": from 0 to 30% long sidings, around 50% long sidings, and above 70% long sidings (Figure 1).

When there are no long passing sidings on the mainline, increasing the percent of long trains increases the average train delay substantially, from 20 min per 100 train-miles (160 train-kilometers) to

over 90 min per 100 train-miles for 75% long trains. As short sidings are extended and the percent long sidings increase, average train delay decreases exponentially through the remainder of the first zone until 50% of the sidings are lengthened. Throughout this zone, scenarios with the highest percent long trains have the largest average train delay but also exhibit the largest decrease in train delay when passing sidings are extended. The delay-based return on investment in extended passing sidings is highest when a greater number of long trains are operated.

In the third zone, where more than 50% of mainline sidings have been lengthened, as the percent long trains increases, overall train delay decreases and a greater reduction in average train delay below the baseline is observed. The results also indicate a smaller relative benefit for additional passing siding extensions.

The trends for all percent long trains in Figure 1 converge to the baseline train delay when 50% of sidings are lengthened. This "transition point" suggests that, for this combination of traffic volume and train lengths, to operate over-length trains at the baseline train delay, only 50% of the mainline sidings need to be extended. Beyond this transition point, long-train operations provide economies of scale and a lower overall train count that reduces delay.

Similar trends are observed for the 2:1, 3:1, and 6:5 length ratios (Figures 2 to 4). Based on these results, when over-length trains are operated on single-track mainlines with short sidings, the train delay response indicates there are two general types of operating behavior separated by a transition point (Figure 5).

Under Type I behavior, the additional delay created by the inflexibility of over-length train meets on mainlines with few long sidings outweighs any delay benefits from reducing the overall train count with longer trains. The outcome is higher average train delay compared to the baseline with all short trains despite the baseline having a higher total train count.

Under Type II behavior, the number of long sidings provides sufficient flexibility in over-length train meets to realize benefits from reduced train count. The mainline operates with less average train delay compared to the baseline condition even though there are still some limitations on where meets between two long trains can be arranged. Though some over-length trains may still be delayed for meets at extended passing sidings, the majority of trains see positive delay benefits arising from the reduced total train count.

The "transition point" between Type I and II behavior is the number of passing siding extensions required to restore the baseline level of service (as measured by average train delay) and mitigate the negative delay effects of over-length train operations. The range of train length ratios in the experiment design was selected to investigate the influence of short and long train lengths on this transition point. Examining changes in the location of the transition point reveals a relationship between the ratio of train lengths and the amount of additional infrastructure required to maintain the initial train delay level of service under long train operations.

When comparing the 3:2 length ratio (Figure 1) to the 2:1 length ratio (Figure 2), the transition point, originally near 50% long sidings for the 3:2 ratio is shifted to the left near 30% long sidings for the 2:1 ratio. The transition point is not constant but changes with train length ratio. As the train length ratio increases, the leftward shift of the transition point suggests that a greater disparity in long and short train lengths requires fewer siding extensions to transition from Type I to Type II behavior. When implementing longer trains at a 2:1 length ratio, only 30% of the mainline passing sidings must be extended to accommodate the longer trains.

In addition to the shift in the transition point location, to the right of the transition point, the lines exhibiting Type II behavior for the 2:1 ratio (Figure 1) are spaced farther apart than those for
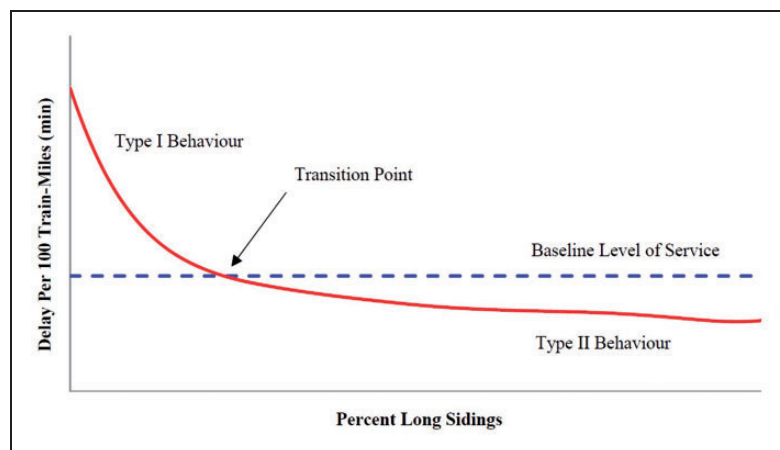

Figure 5. Average train delay behavior of over-length train operations under increasing number of passing siding extensions.

the 3:2 ratio (Figure 2). Where long sidings are more frequent, the reduction in delay obtained from operating long freight trains increases with the train length ratio. Higher train length ratios result in lower overall train counts, fewer train meets, and less delay.

The other two train length ratios (Figures 3 and 4) support the relationship between the transition point and train length ratio. For the highest length ratio (3:1), the transition point moves to its lowest value of extended passing sidings. For the lowest length ratio (6:5), the transition point moves to its highest value of extended passing sidings. There appears to be a linear relationship between required infrastructure investment at the transition point (measured in percent long sidings) and train length ratio, with the transition point decreasing as length ratio increases (Figure 6).

Although inferring a linear relationship from only four data points must be undertaken with care (see below for further comments), the form of the relationship between the transition point and train length ratio has potential applications as railway industry practitioners consider siding extension programs and optimizing train length. A railway infrastructure owner that desires long train operations may only have sufficient budget to lengthen a certain percent of their passing sidings. Assuming similar route and traffic characteristics, this owner can use the derived relationship to estimate what train length is required to maintain their current level of service (as measured by average train delay) given their budgeted siding extensions. Alternatively, if a desired train length has been defined along with its corresponding length ratio, an estimate of the required number of siding extensions can be developed from Figure 6. As a linear relationship, a unit increase in the length of long trains consistently translates to the same unit decrease in the required number of passing siding extensions. Given the factors considered in this
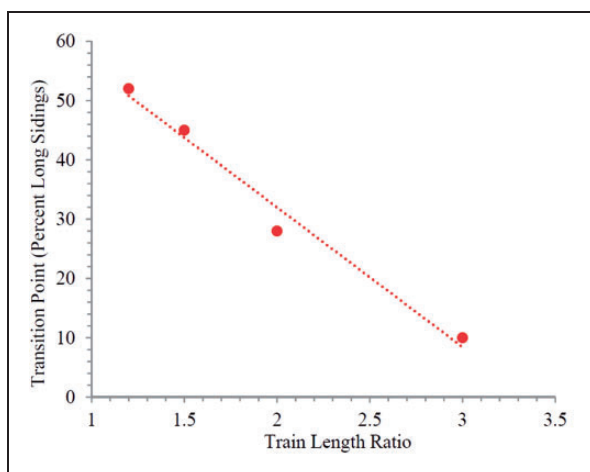
study and ignoring other operating and terminal costs, there does not appear to be an optimal train length ratio or point of diminishing returns where further increases in train length are not accompanied by additional reductions in the required number of extended passing sidings.

The linear nature of the relationship can be argued from a conceptual standpoint. For a train length ratio of 1.01 (101-car long trains, 100-car short trains), the introduction of long trains would do little to reduce the overall train count. Without any reduction in delay from a lesser train count, it can be hypothesized that nearly all of the mainline sidings must be extended to maintain the current level of service. As the length ratio approaches unity, the transition point could be expected to approach 100% long sidings. A hypothetical length ratio of 12 (1200-car long trains, 100-car short trains) may not require any passing sidings to be extended since operating two 1200-car trains (one in each direction) can achieve a 2400-car throughput. The transition point may asymptote to zero as the train length ratio takes on large values. These conceptual points would not fall on the linear relationship presented in Figure 6, suggesting it may only be linear over the range of values tested and then curving to asymptote the axes in the form of an inverse function. Simulation of additional scenarios at these extremes could confirm these hypotheses and the overall shape of the relationship. However, the most common over-length train scenarios considered by the North American rail industry fall along the locally linear result. The linear assumption can aid decisions on siding extension and train lengthening programs across the range of operating conditions encountered in the North American railway industry.

Relative delay to short and long trains

The main experiment design only considered average train delay and the bidirectional operating strategy. To better understand the relative behavior of short and long trains, additional simulations were conducted for both the bidirectional case and the unidirectional case with all long trains operating in the same direction. The additional simulations were conducted at a traffic volume of 3600 railcars per day, 50% long trains, and 3:2 length ratio (150-car long trains and 100-car short trains). Following completion of all simulation experiments, train delays for long and short trains were separately totaled and normalized. For comparison purposes, scenarios with 36 short (100-car) trains were also simulated across the range of percent long sidings to serve as a train delay baseline.

When the short and long trains are operated in a bidirectional manner and there are few long sidings (Type I behavior), both short and long trains exhibit increased train delay relative to the baseline

Figure 6. Train length ratio and the percent long sidings at the transition point.

(Figure 7). Long train delays are up to four times higher than the baseline train delay. If a corridor does not have long sidings, dispatchers cannot simultaneously grant movement authority to long trains in both directions; long trains must wait at end terminals and accumulate additional delay. By extending a single short siding, long trains can meet along the route and no longer need to spend as much time waiting in terminals, reducing delay by 25%. Most of the conflicts between short and long trains are handled in favor of long trains, where a long train proceeds without stopping and a short train stops on the passing siding.

With the introduction of long trains and long–short train conflicts, the average delay of short trains rises from 40 to 80 min per 100 train-miles. This increase in delay is attributed to long–short train conflicts where short trains are diverted into the passing sidings where a longer train cannot fit. Extending passing sidings helps reduce the delay of short trains in an indirect way; by improving the

resolution of conflicts between long trains, the long trains traverse the route in less time and meet fewer short trains.

As the level of infrastructure expansion approaches the percent long sidings at the transition point, the difference in train delay between long and short trains decreases. When a sufficient number of passing sidings are extended, conflicts between any two trains can be resolved in an efficient manner regardless of length, and short and long trains accumulate delay at the same rate. For Type II behavior, both short and long trains benefit from the overall reduction in train count and exhibit less train delay than the baseline.

For the unidirectional case where all long trains operate in the same direction and there are no meets between two over-length trains, a different pattern of train delays is observed (Figure 8). When there are few long sidings on the corridor, long trains exhibit low train delays while short trains experience an increase in average train delay relative to the baseline. With few long sidings, the dispatching logic
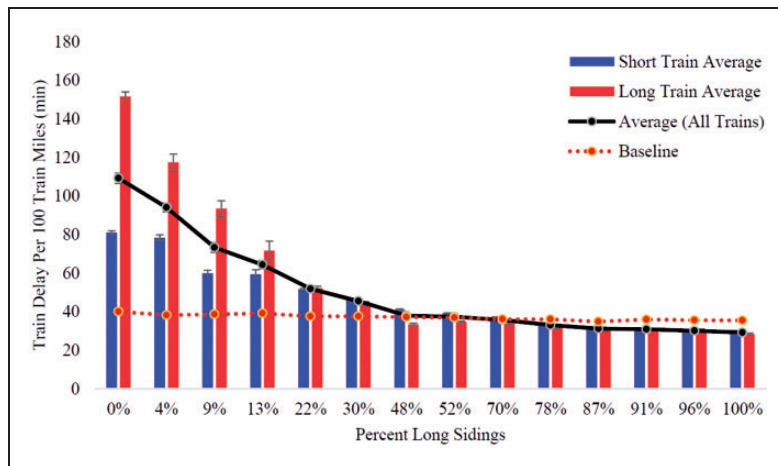

Figure 7. Train-type delay for a range of percent long sidings and 50% bidirectional long trains, 3:2 length ratio, 3600 railcars per day.
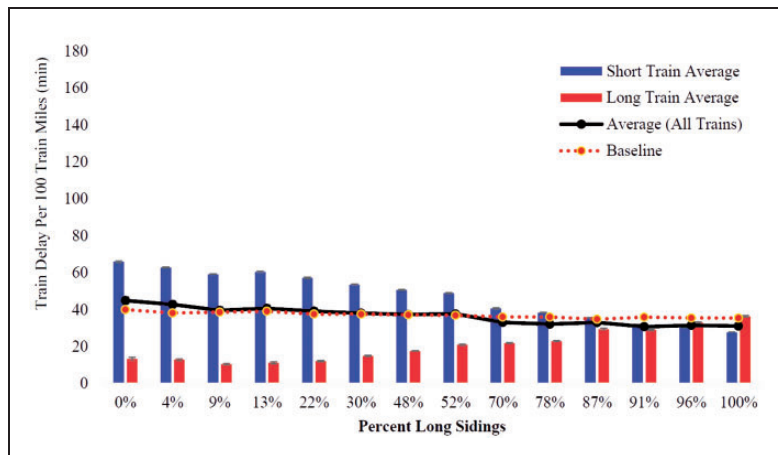

Figure 8. Train-type delay for a range of percent long sidings and 50% unidirectional long trains, 3:2 length ratio, 3600 railcars per day.

consistently routes short trains into passing sidings to wait and accumulate delay while long trains are effectively "prioritized" and rarely stop. The long trains proceed uninterrupted to the destination terminal and with low delay. When the number of long sidings on the line increases, conflicts can occur at sidings that can handle both long and short trains. Both train types become interchangeable to a dispatcher and train-type delay values equalize; short trains benefit from passing siding extensions while long trains actually experience a counterintuitive increase in delay.

Practitioners that are particularly concerned with the delay performance of long trains on routes with few long sidings should consider unidirectional operation in the direction with the largest quantity of long trains. If possible, given the network layout, opposing long trains should potentially be rerouted or split into shorter trains to avoid delay-intensive conflicts between long trains. However, the delay benefits of unidirectional long train operations must be weighed against potential additional complication, delay, and operating expense at terminals where unit trains are lengthened or shortened before returning in the opposite direction. Also, the short and long trains may have different business objectives and tolerances for delay, further influencing the decision between bidirectional and unidirectional operations of long trains.

To further expand on the previous discussion, the data in the supplemental simulations were reanalyzed for several percent-long-siding infrastructure configurations to determine the average number of stops and average delay per stop for each train type (Figure 9(a) to (d)). The average number of stops by train type is plotted on the horizontal axis and average delay per stop is plotted on the vertical axis. The results described below are on a total of 40 simulated cases. The data points acquired for each are based on five simulation days, 10 replications, and 30 trains. The total number of stops is calculated by aggregating all stops over a five-day period. Average number of stops is calculated by dividing the total number of stops by five. Average delay per stop is calculated by dividing the total train delay by the total number of stops over a five-day period. Each data point is replication and train specific.

The changing distributions of average number of stops and delay per stop as percent long sidings increase help explain the operational mechanics behind the train-type delay responses in Figures 7 and 8.

When no passing sidings are extended (Figure 9(a)), long trains only dwell at terminals to wait for opposing trains to arrive. These dwells can be lengthy as a train may need to wait for an opposing long train to traverse the entire route before it can proceed. Thus, long trains exhibit a low number of stops but with a high average delay per stop.

As the number of siding extensions starts to increase (Figure 9(b) and (c)), long trains can dwell to meet opposing long trains at the long sidings. Due to the limited number of long sidings, arranging
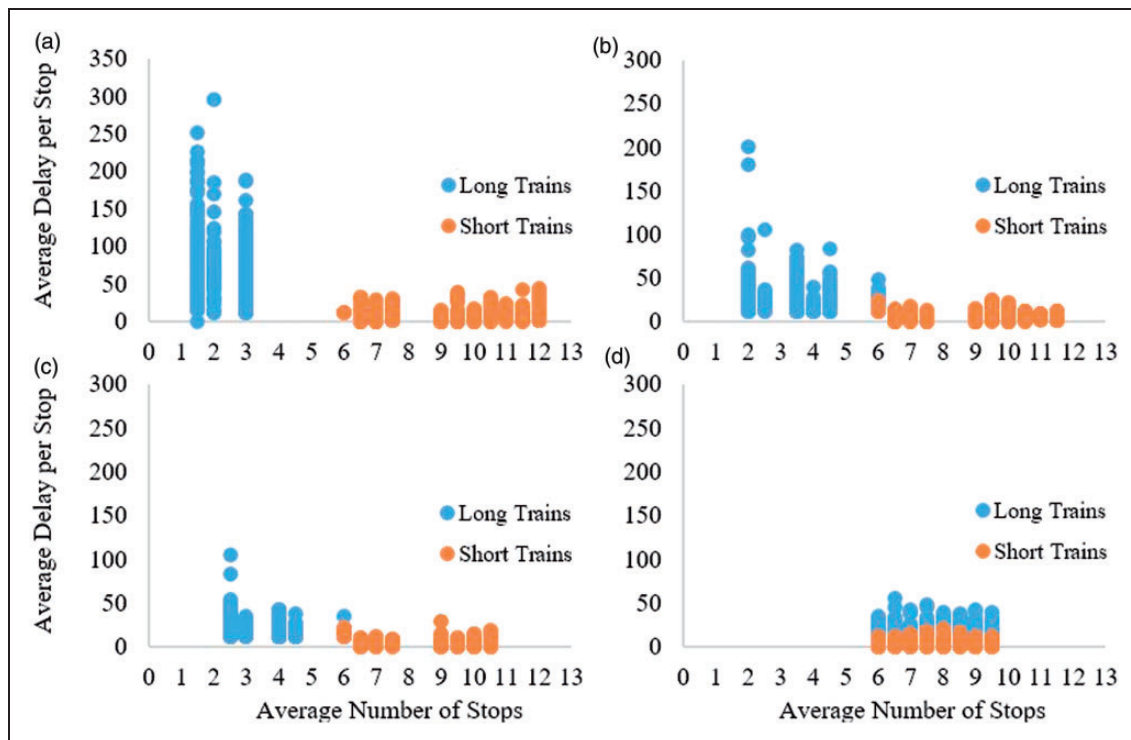

Figure 9. Number of stops and train-type delay per stop for bidirectional operation, 3:2 length ratio, 3600 railcars per day and (a) 0%, (b) 22%, (c) 52%, (d) 100% long sidings.

a meet between opposing long trains imposes a high delay as one long train will typically experience an extended dwell for the second train to arrive. Long trains feature less frequent but more prolonged stops when compared to the short train conflict distribution.

When all passing sidings on the corridor are extended, there is little operational distinction between long and short trains and both train types display a similar average number of train stops and train delay per stop (Figure 9(d)). When there are many long sidings, the dispatching logic sees merit in routing a long train into a passing siding during a conflict with a short train. Consequently, the average number of stops for short trains decreases while the average number of stops for long trains increases. With neither train type being favored, the average delay per stop becomes equivalent.

Conclusions

In North America, where the majority of mainlines are single track, the potential operational and economic advantages of long trains are limited by the inadequate length of many existing passing sidings. To resolve conflicts between over-length trains, mainline track infrastructure must be expanded by extending existing passing sidings or constructing new longer passing sidings. The presented research demonstrates that the number of passing siding extensions required to maintain a given level of service (quantified by train delay) is related to the train replacement ratio and train length ratio that describes the relative lengths of the long short trains operating on a particular mainline. Simulation results indicate there is a transition point where enough existing passing sidings are extended to allow flexibility in long train meet locations and restore the original level of service after the introduction of over-length trains. There is a declining linear relationship between the train length ratio and this transition point as quantified by the required number of passing siding extensions. Fewer passing siding extensions are required to obtain the economies of long-train operations when the train length ratio is large. Practitioners engaged in railway infrastructure and operations planning can use these findings to evaluate return on investment and help establish the magnitude of programs to extend existing passing sidings in support of longer freight trains.

With bidirectional operation of over-length trains, there are distinct differences in the train delay performance of long and short trains: long trains experience approximately twice as much delay as short trains when there are few extended passing sidings. Efforts to reduce train delay and improve levels of service through passing siding extensions show less return for short trains in comparison to long trains. With unidirectional operation of over-length trains, long trains experience delay values far below the baseline level of service at low levels of infrastructure investment. Knowledge of the relative delay performance of different length trains can influence practitioner decisions on the appropriate size of trains to meet the business objectives and reliability goals of different freight train services.

ORCID iD

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

References

1. Ede BM, Polivka A and Tunna J. Optimizing capacity by trade-offs among train control, infrastructure, mechanical, and engineering considerations. In: *Proceedings of the 2007 IHHA specialist technical session*, Kiruna, Sweden, 2007, pp.413–421. International Heavy Haul Association.
2. Barrington B and Peltz D. 10,000-ft distributed power intermodal trains. In: *Proceedings of the ninth International Heavy Haul Association conference*, Shanghai, China, 2009.
3. Martland CD. Introduction of heavy axle loads by the North American rail industry, 1990 TO 2012, In: *Proceedings of the 54th annual transportation research forum*, Annapolis, Maryland, 2013, pp.36–53. Transportation Research Forum.
4. Starr JT. *The evolution of the unit train, 1960–1969*. Chicago, IL: University of Chicago, 1976.
5. Beck R, Connell D and Iden M. Western coal service on union pacific railroad: operations and maintenance update. In: *Proceedings of the international heavy haul association specialist technical session*, Dallas, Texas, 2003.
6. Association of American Railroads. *Railroad facts*. 2012 ed. Washington, DC: Association of American Railroads, 2012.
7. Bell S. Driving efficiency, preserving infrastructure and creating capacity with long trains. In: *Proceedings of the 10th International Heavy Haul Association conference*, New Delhi, India, 2013, pp.789–795. International Heavy Haul Association.

8. Keaton MH. Service-cost tradeoffs for carload freight traffic in the U.S. rail industry. *Transp Res Part A* 1991; 25: 363–374.
9. Richards CW and Cobb T. Multiple-track main lines. *Trains*, January 2006.
10. Sogin S, Lai YC, Dick CT, et al. Analyzing the transition from single- to double-track railway lines with nonlinear regression analysis. *J Rail Rapid Transit* 2016; 230: 1877–1889.
11. Blaine DG, Hengel MF and Peterson JH. Train brake and track capacity requirements for the '80s. *J Eng Ind Trans Am Soc Mech Eng* 1981; 103: 392–401.
12. Welch N and Gussow J. Expansion of Canadian National Railway's line capacity. *Interfaces* 1986; 16: 51–64.
13. Grimes GA. Planning for unitrain operations on Missouri Pacific. In: *Proceedings of the transportation research forum*, San Francisco, California, 1981, pp.121–126. Transportation Research Forum.
14. Newman RR, Zarembski AM and Resor RR. Economic Implications of heavy axle loads on equipment design operations and maintenance. In: *American Society of Mechanical Engineers*, WAM RTD-Volume 4, Rail Transportation, December 1991.
15. Barton R and McWha T. Reducing emissions in the rail sector: technology and infrastructure scan and analysis. Report, Transport Canada Transportation Development Centre, Canada, 2012.
16. Martland CD. Railroad train delay and network reliability. Report R-991, Association of American Railroads, Transportation Technology Center, Inc., USA, 2008.
17. Dick CT and Clayton AM. Impact of new railway technology on grain transportation in Western Canada. *Transp Res Rec J Transp Res Board* 2001; 1742: 45–53.
18. Atanassov I and Dick CT. Capacity of single-track railway lines with short sidings to support operation of long freight trains. *Transp Res Rec J Transp Res Board* 2015; 2475: 95–101.
19. Atanassov I. *Influence of track arrangement on expanding rail corridor capacity and operations.* Master's Thesis. University of Illinois at Urbana-Champaign, Department of Civil and Environmental Engineering, USA.
20. Krueger H. Parametric modeling in rail capacity planning. In: *Proceedings of the 1999 winter simulation conference*, Institute of Electrical and Electronics Engineers (IEEE), Phoenix, Arizona, 1999, pp.1194–1200. IEEE.
21. Dick CT and Mussanov D. Operational schedule flexibility and infrastructure investment: capacity trade-off on single-track railways. *Transp Res Rec J Transp Res Board* 2016; 2546: 1–8.
22. Pouryousef H, Lautala P and White T. Railroad capacity tools and methodologies in the U.S. and Europe. *J Mod Transp* 2015; 23: 30–42.
23. Berkeley Simulation. *Rail Traffic Controller (RTC) User Manual*. Berkeley, California: Berkeley Simulation Software, July 2011.
24. Sogin S, Lai YC, Dick CT, et al. Comparison of capacity of single- and double-track rail lines. *Transp Res Rec J Transp Res Board* 2013; 2374: 111–118.
25. Shih MC, Dick CT and Barkan CPL. A parametric model of the train delay distribution based on traffic conflicts. In: *Proceedings of the International Association of Railway Operations Research (IAROR) 7th international conference on railway operations modelling and analysis*, Lille, France, 2017.
26. Sehitoglu T, Mussanov D and Dick CT. Operational schedule flexibility, train velocity and the performance reliability of single-track railways. In: *Proceedings of transportation research board 97th annual conference*, Washington, DC, USA, 2018.
27. Atanassov I and Dick CT. Influence of siding connection length, position, and order on the incremental capacity of transitioning from single to double track. In: *Proceedings of transportation research board 94th annual conference*, Washington, DC, USA, 2015.