INFLUENCE OF TRACK ARRANGEMENT ON EXPANDING
RAIL CORRIDOR CAPACITY AND OPERATIONS

BY

IVAN H. ATANASSOV

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Advisers:

Professor Christopher P.L. Barkan
Senior Research Engineer C. Tyler Dick, P.E.

ABSTRACT

The North American freight railroad network is projected to experience rising freight transportation demand in the coming decades, coupled with continued interest in expanding passenger services. Congestion resulting from these demands strains the capacity of rail lines and jeopardizes the operational fluidity of the rail network, particularly along shared rail corridors. While track construction is just one of many alternatives a railroad may employ in expanding practical capacity (and thereby boosting throughput), this practice represents substantial capital investment. With the purpose of helping rail practitioners better utilize their resources, this thesis aims to investigate track expansion alternatives in detail, ultimately providing an improved understanding of the link between track arrangement, train delay, and line capacity.

The majority of mainline rail corridors in North America consist predominantly of single track with passing sidings or short sections of double track. These track arrangements lack the flexibility to reliably handle high traffic volumes composed of multiple types of trains. Increasing frequency of long freight-train operations also magnifies capacity constraints posed by inadequate, short sidings. This work explores the capacity benefits of siding expansion to meet these developing operational needs, leading to a discussion of the incremental capacity in transitioning from single to double and triple track, both from a quantitative and qualitative perspective. Experiment designs are carried out in Rail Traffic Controller simulation software to reveal fundamental relationships between track arrangement and other capacity factors via statistical analysis of the results. While railroads must consider many factors in selecting capital expansion projects, the trends identified through this research can help streamline the planning process by helping industry practitioners quickly identify track expansion project alternatives with the greatest potential capacity benefit for more detailed engineering evaluation.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 Purpose

This work aims to quantify relational trends between railway track arrangement and line capacity, ultimately helping rail practitioners to efficiently use their planning resources by providing initial screening for costly track expansion project alternatives.

1.2 Background

The operational landscape of North American railways is currently undergoing changes that will continue into the coming decades, due in part to changing markets. In the United States, these changes can be attributed to a forecasted rise in demand for freight rail transportation that, coupled with increased interest in providing new, faster, and more frequent passenger rail services, will lead to increased congestion along many rail lines (Association of American Railroads 2008). The simultaneous increase in demand for freight and passenger rail transportation will place particular strain on existing shared rail corridors where freight and passenger trains operate on the same track infrastructure.

The majority of mainline corridors in North America consist predominantly of single track with passing sidings or short sections of double track. These track arrangements, while adequate for moderate traffic volumes, lack the flexibility to handle high traffic volumes composed of multiple types of trains operating over a wide range of speeds. Increasing frequency of long freight-train operations also amplify capacity constraints posed by inadequate, short sidings on many single-track routes. A study conducted by Cambridge Systematics (2007) estimated that without improvements to the national rail infrastructure in the United States, thirty percent of rail miles along primary corridors would operate above available capacity given projected traffic volumes in the year 2035. This condition is reached without any future modal

shift of freight from truck to rail, nor any expansion of passenger service. Consequently, continued railway industry investment of private capital in additional track infrastructure will be necessary, and selection of projects that provide the highest returns is imperative. Conventional evaluation of possible track expansion alternatives via detailed rail simulation software platforms, however, requires substantial use of time and resources. Greater knowledge and understanding of fundamental relationships between track arrangements and rail line capacity can help railroads more efficiently use network planning and engineering resources by improving the initial alternative selection process and reducing the number of alternatives designated for detailed simulation analysis. The work presented here investigates the link between these fundamental relationships by conducting simulation experiments designed to quantify capacity-related benefits afforded by siding expansion projects, as well as incremental transitions from single to double and triple track.

1.3 Objective and Scope

This thesis considers a range of track and train characteristics typical of North American mainline operations. Nevertheless, the simulated rail corridors are somewhat idealized from infrastructure and operating perspectives (e.g. balanced track distribution, uniform speeds, etc.). The experiments simplify real-world infrastructure and operating conditions to reduce sources of variation and allow for a focus on the fundamental relationships between key variables of interest (i.e. train delay, traffic volumes, etc.). The research presented here is also based on mainlines that are simulated as isolated entities, rather than part of the larger network and with links to adjacent rail lines through terminals. Terminals and network effects have a direct impact on mainline capacity, albeit one that has not yet been well defined, and is therefore a point of future work in subsequent chapters.

The use of simulation for conducting the experiments in this thesis has practical limitations. The Rail Traffic Controller (RTC) simulation software used in this work (described in the following chapter) is robust and takes into consideration numerous factors, but it cannot consider every nuance and unplanned event in railway operating environments. Furthermore, the simulation software only *emulates* real-time dispatching, rather than mimicking it. Specific values obtained via simulation are more aptly considered as relative values rather than absolutes. Comparing one set of simulation results against another helps uncover relationships among variables, and is thus useful for the purpose of this work – quantifying relational trends between track arrangement and capacity.

1.4 Organization

This thesis is composed of eight chapters that, following this introduction, document various components of the research as described in the subsequent paragraphs.

CHAPTER 2 – DETERMINING THE PRACTICAL CAPACITY OF DIFFERENT
TRACK ARRANGEMENTS VIA SIMULATION

Existing literature on railroad capacity-related topics is presented in this chapter, including an explanation of the logical progression from prior studies to those discussed in the following chapters. Not only are results from previous works discussed, but the tools and methods used to obtain the results are also introduced. The process used to build and specify cases in RTC (e.g. file creation, working with infrastructure, parameter specification, etc.) is also referenced in this chapter.

CHAPTER 3 – LONG-TRAIN OPERATION WITH SHORT SIDINGS

In North America, the majority of mainline routes are single track with passing sidings. The potential economic and operational advantages offered by long freight trains are constrained

by the inadequate length of many of these sidings. Chapter 3 analyzes train delay resulting from operating trains whose lengths exceed the longest sidings on a corridor, and discusses practical implications of the results. An earlier version of this research has been accepted for publication in the Transportation Research Record: Journal of the Transportation Research Board.

CHAPTER 4 – LONG-TRAIN REPLACEMENT RATIO AND INFRASTRUCTURE INVESTMENT

This chapter extends Chapter 3 by examining different combinations of short and long trains and introducing the concept of "train replacement ratio". This study seeks to more formally understand the level of infrastructure investment (i.e. number of longer passing sidings) required before the operational efficiencies afforded by running longer trains are realized in the context of train delay. An earlier version of this research appears in the proceedings of the 2015 International Heavy Haul Association (IHHA) conference in Perth, Australia.

CHAPTER 5 – SINGLE TO DOUBLE TRACK: INFLUENCE OF SIDING SPACING

The discussion in this chapter is based on prior studies of the train-delay benefits of incrementally adding segments of double track between passing sidings on single-track mainlines. Prior research is expanded on by investigating the capacity effects of double-tracking routes with relatively longer distances between passing sidings, spaced evenly throughout the route. Given that even siding spacing is difficult to achieve in practice due to various engineering constraints, a more realistic scenario is introduced to quantify train delay in response to double-tracking a route with different combinations of distances between sidings. The studies conducted in this chapter are focused on homogeneous freight traffic. An earlier version of this research appears in the proceedings of the 2014 Joint Rail Conference in Colorado Springs, USA.

CHAPTER 6 – SINGLE TO DOUBLE TRACK: SIDING CONNECTION
LENGTH, POSITION, AND ORDER

This chapter is an extension of Chapter 5 and introduces heterogeneous (i.e. mixed passenger and freight) traffic into the simulation experiment design. This experiment attempts to determine the relative influence of siding spacing, position on the route, and connection order on incremental capacity of double-track segments. An earlier version of this research appears in the proceedings of the Transportation Research Board 94th Annual Meeting in Washington, D.C. in January 2015.

CHAPTER 7 – INCREMENTAL CAPACITY IN TRANSITIONING
FROM DOUBLE TO TRIPLE TRACK

Full double track has finite capacity under high volumes of mixed rail traffic (e.g. speed and train-size heterogeneity). To allow for higher traffic volumes and complex operating patterns, additional main tracks may become necessary. Thus, the incremental capacity in transitioning from two to three main tracks is quantified by simulation experiments of heterogeneous operations. Consideration is also given to different crossover configurations and their effect on line capacity. An earlier version of this research appears in the proceedings of the International Association of Railway Operations Research (IAROR) 6th International Seminar on Railway Operations Modelling and Analysis in Tokyo, Japan in March 2015.

CHAPTER 8 – CONCLUSIONS AND FUTURE WORK

This chapter provides a summary discussion of the combined results and their implications for railway operators and planners. Suggestions for railway capital improvement studies are outlined to introduce a more streamlined evaluation process for infrastructure alternatives. Suggestions for future experimentation and analyses are also presented.

CHAPTER 2
DETERMINING THE PRACTICAL CAPACITY OF DIFFERENT TRACK ARRANGEMENTS VIA SIMULATION

2.1 Literature Review

2.1.1 Defining Capacity

Before hypotheses relating track arrangement to rail line capacity can be addressed, the term 'capacity' must be defined in a railway context. Capacity does not conform to any one strict definition; rather, the term takes on a number of interpretations that depend on its application and use to achieve specific goals. For example, in a report prepared for the Association of American Railroads by Cambridge Systematics (2007), capacity is associated with the volume of freight (in tons) moving across a particular line. This same definition can be translated to passenger systems, where capacity can be measured by the number of persons being moved across a line during a given time period. This interpretation of capacity, while easily communicated to the public and stakeholders, refers to the overall throughput of a rail line, but does not give any indication of the provided level of service and reliability of the operation. Thus throughput metrics of capacity can be deceptive when comparing lines with different types of trains, each with their own service requirements.

To address this shortcoming, in subsequent sections of this thesis, capacity is measured in terms of train delay (i.e. units of time per train). More specifically, train delay serves as a metric for capacity by comparing actual train run times between terminals, or Total Elapsed Time, against the scheduled, ideal run times, or Ideal Run Time (Equation 2.1).

$$Delay = Total\ Elapsed\ Time - Ideal\ Run\ Time \qquad\qquad (2.1)$$

Ideal Run Time is the time required for a train to travel between origin and destination while making all planned service stops but without any interference from other trains. In subtracting ideal run time from actual run time, the resulting delay takes into consideration the time spent stopped on the mainline, proceeding at slower speeds, and acceleration/deceleration resulting from train conflicts.

Since delay accumulates over the length of a train run, delay is normalized per 100 train-miles to allow for comparisons between routes and train-runs of different lengths. Delay serves as both a metric of level of service and a proxy for line capacity in Kreuger's (1999) work using delay-volume curves. The characteristic delay-volume curve for a route allows delay to be related to a maximum traffic throughput corresponding to that level of service. This method for defining and communicating capacity has appeared more recently in the academic works of Sogin et al. (2013a) and Dingler et al. (2013) that addressed capacity of rail lines with different track arrangements.

While train delay is less easily interpreted than pure train count, using delay as a metric for capacity produces finer-grained analyses that are useful to both freight and passenger rail operators alike. For example, a capacity study for a freight operator may find that a particular line can theoretically handle two more trains per day if capacity is solely defined via train throughput. However, adding two extra trains, while technically feasible, may dramatically increase train delay on the route such that levels of service degrade to undesirable levels, and the route is considered oversaturated. As for passenger services, train delay may be a natural choice for capacity analysis given the scheduled nature of operations. A hybrid definition of capacity that gives consideration to both throughput and train delay can also be used (Lindfeldt 2006).

2.1.2 Track Arrangement and Capacity

Literature pertaining to the relationships between track arrangement and line capacity is extensive, reflecting the breadth of a practical topic with identifiable impacts on railway capital planning and expenditures. As outlined earlier, however, the subsequent sections in this work focus on three main research areas: the capacity associated with long-train operations on routes with inadequate siding lengths, the incremental capacity in transitioning from single to double track, and the transition from double to triple track – topics less studied in the context of North American freight and shared-corridor operations.

The need to investigate siding extension programs to facilitate operation of longer freight trains is put in context by Martland (2013), who commented on the insufficiency of existing passing sidings to handle long-train operations on single track. Martland observed that two-thirds of unit trains operating in the United States are "length-limited" by passing sidings, and that this estimate was conservative. Jaumard et al. (2013) used a dynamic management algorithm and optimization model to simulate long-train-short-train interactions along a shared line. Their study indicated that in order to successfully incorporate longer-train operations along a route, departure-time scheduling must be considered in a joint process. Kraft (1982) also used analytical tools to discuss fleeting techniques for long trains and to analyze the capacity benefit of running longer trains on a representative route with a mixture of short and long sidings.

In regards to investigating the incremental capacity in transitioning from single to double track, the work presented in the following chapters is inspired by that of Sogin et al. (2013a). This research identified relationships between train delay and varying levels of double track, and ultimately created response surface models in the form of delay-volume curves. Results revealed that for idealized single-track corridors with evenly-spaced passing sidings, double-track

installation provided a linear reduction in freight train delay when traffic volume was held constant. The benefits of double-track segments are also considered by Lindfeldt (2012), who notes improved timetable flexibility from the addition of double track that in turn imposes a higher practical, realizable capacity.

As mentioned earlier, high traffic volumes composed of multiple types of trains operating at different speeds while sharing infrastructure could make triple track a viable alternative for achieving fluid operations. In the study by Cambridge Systematics (2007), the practical capacity of double track dropped from 100 to 75 trains per day once heterogeneous operations were introduced. This indicates the potential need for triple track to alleviate congestion resulting from train counts above this threshold. Tobias et al. (2010) used simulation models to investigate the inability of double track to provide sufficient capacity to sustain the expected 20-year passenger and freight traffic growth along a particular shared-use rail corridor in the United States, and forecasted the physical need for triple-track installation to remedy these operational maladies.

While the literature referenced above provides a brief outline of methods and results that have appeared in studies related to the topic of this thesis, more detailed discussion of previous work on specific sub-topics can be found in respective chapters.

2.2 Rail Traffic Controller

The experiments presented in the following chapters develop capacity (train delay) metrics using Rail Traffic Controller (RTC), the *defacto* industry-standard rail traffic simulation software in the United States. Specially developed for the North American railway operating environment, RTC emulates dispatcher decisions in simulating the movement of trains over rail lines subject to specific route characteristics (Wilson 2015). RTC is used extensively by a wide range of public and private organizations, including most Class I railroads, Amtrak, Bay Area

Rapid Transit (BART), and major railroad consultants. Inputs for simulations run in RTC include factors such as track arrangement, signaling, speed limits and train consists. Outputs include, but are not limited to, reports on train delay, dwell, siding usage, and train energy consumption.

An RTC methodology for rail capacity studies was documented by Sogin (2013), whose software implementation and conventions served as the basis for the research described here. For the work presented in subsequent chapters, RTC inputs are varied to reflect changes in track arrangements and train parameters, with train delay as the output. To determine a train delay response, each unique combination of input variables (including track arrangement) in an experiment design is simulated in RTC for five days of rail traffic. To allow for variation in train departure times, each simulation is replicated five times to provide 25 days of train operations used in calculating average train delay. To be consistent with flexible North American freight rail operating practices, each replication uses a different train operating pattern where the specified number of trains per day (traffic volume) depart at random intervals from their respective terminals during a 24-hour window. The random train departures are generated from a uniform distribution over each 24-hour period; they are not distributed around a particular target departure time. Thus, there is no pre-determined departure and arrival schedule and the locations of meets and passes between trains are not pre-established by an operating timetable.

While this approach to train scheduling and replication applies to all simulations completed for this thesis, more comprehensive RTC methodologies tailored to each study are presented in respective chapters.

CHAPTER 3
LONG-TRAIN OPERATION WITH SHORT SIDINGS

3.1 Introduction

Increasing the length of freight trains provides economies of scale with respect to fuel consumption and operating crew costs, and positively affects line capacity by reducing the number of trains required to move a given freight volume (Moore et al. 2007, Barrington & Peltz 2009). In 1980, the average freight train in the western United States contained 68.9 railcars, but by 2000 this had only increased to 72.5 railcars. Over the past decade, increasing use of distributed power and AC-traction locomotives in North American heavy-haul service has allowed for greater efficiencies through regular operation of freight trains in excess of 125 railcars in length. In 2010, the average train had grown to 81.5 railcars and railroads had begun to operate 150-car trains on selected corridors (Association of American Railroads 2012). Thus, longer freight trains are still a relatively new phenomenon in the North American rail industry.

The implementation of long freight trains on existing routes is contingent on the physical capacity of the existing route infrastructure to handle these longer trains. The railway infrastructure in North America is primarily composed of single-track mainlines with passing sidings whose lengths were sized for 100-car trains prevalent at the time of construction. These passing sidings are inadequate for staging meets between two new, longer freight trains. Meets between two long trains must be carefully planned to occur at extended-length sidings, on sections of double track, or within terminals with adequate track capacity. This operating constraint reduces flexibility and potentially introduces congestion and delay that may partially

offset the efficiencies afforded by long trains. As a result, freight infrastructure owners must adopt capital expansion programs that focus on the extension of existing passing sidings, or the construction of new longer-length passing sidings, to provide the physical capacity required to serve longer freight trains.

The analyses that follow aim to characterize the relationship between the lengths of single-track rail corridor passing sidings and the operation of long freight trains from the perspective of line capacity (as measured by train delay). This research considers the problem of mismatched siding and train length by using archetypal infrastructure and train characteristics in in an experiment to quantify the relationship between the number of long sidings and the practical number of long trains that can operate on a route. While there are many factors to consider in the planning stages of rail infrastructure expansion, the results of this study can streamline the planning process by establishing general guidelines for the number and types of passing-siding extension and construction projects with the highest expected return on investment.

3.2 Background

Interest in operating long freight trains in heavy-haul service, as well as their economical and operational efficiency, has been well documented in the literature, from both a numerical and qualitative perspective. Newman et al. (1991) described the economic and operational benefits of increasing the length of unit trains on one Class I railroad. Operational advantages of longer freight trains are discussed by Barton and McWha (2012), who cited the need for lengthened passing sidings in response to freight trains up to 12,000 feet in length by several North American Class I railroads. The sentiment for siding extension programs was shared by Martland (2013), who elaborated on the insufficiency of existing passing sidings to handle long-train

operations by his conservative estimate that two-thirds of unit trains operating are "length-limited" by passing sidings. The ability of siding length to dictate the maximum practical length of trains on a particular corridor was also discussed by Dick and Clayton (2001), who demonstrate that, at the time of writing in 2001, most sidings on Canadian Pacific Railway (CP) and Canadian National Railway (CN) were of insufficient length to adequately support long-train operations. To overcome its siding-length disadvantage relative to CP, competitor CN began to run 150-car trains (9,000 feet in length) in a single direction to avoid the problem of meets between long trains. For perspective, typical sidings range in length from 6,000 to 7,500 feet, or from 100 to 125 railcars.

The efficiency of longer freight trains, as well as their interaction with relatively shorter sidings, has been researched from a more analytical perspective by Jaumard et al. (2013). A dynamic management algorithm and optimization model were used for the purpose of simulating long-train-short-train interactions along a shared line. Kraft (1982) also used analytical tools to discuss fleeting techniques for long trains and to analyze the capacity benefit of running longer trains on a representative route with a mixture of short and long sidings.

The research presented in both this chapter and the next aims at expanding upon the aforementioned research on long-train operability to address three key research questions:

- Although it is intuitive that introducing long trains to a route with no long sidings will disrupt operations, what is the exact impact on train delay relative to the required level of service?

- While long-train operations can be supported by extending all passing sidings on a route, this represents a large capital investment. For different mixtures of long

and short trains, can the required level of service be maintained by extending a limited number of passing sidings along a route?

- Is the required number of sidings a function of the number of long trains relative to the total traffic on the route?

The above questions are addressed by conducting a detailed simulation experiment that quantifies the specific relationship between the number of long sidings on a route and the number of long freight trains that can be operated at a given level of service.

3.3 Methodology

This study conducts an experimental design matrix of simulations on a representative route whose general characteristics, along with the properties of the freight trains, are typical of North American freight operations (Table 3.1). The experimental design matrix itself is comprised of four main variable factors: total freight throughput, percent long sidings, percent railcars in long trains, and the directional distribution of long trains operating on the route.

Table 3.1: Simulated route and freight train characteristics

| Route & Train Characteristics	Values
Length	240 miles
Siding Spacing	10 miles
Total Number of Sidings	23
Siding Lengths	2mi (long), 1.25mi (short)
Traffic Composition	100% Freight
Locomotives	SD70 (x2 or x3)
Number of Cars	100 (short train), 150 (long train)
Total Length of Cars	5,500ft (short train), 8,250ft (long train)
Maximum Freight Speed	50mph (45mph through siding)
Traffic Control System	2-block, 3-aspect CTC

- **Total freight throughput** is the number of railcars per day moved across the representative subdivision. To provide a constant level of transportation productivity, the total number of railcars moved during a single day by short and long trains combined must equal the specified total freight throughput.

- **Percent long sidings** is defined as the fraction of total sidings on the route that are longer than the length of long trains; in this case, 150 railcars.

- **Percent railcars in long trains** is the fraction of total railcars moving in long trains. For example, if the baseline traffic of 3,600 railcars per day consists of 36 short trains, each 100 cars in length, the case of 50 percent railcars in long trains consists of 18 short 100-car trains and 12 long 150-car trains.

- **Directional distribution** specifies how many of the long trains are operating in each direction. A 50-50 directional distribution is the bi-directional case with an equal number of long trains in each direction. A 100-0 directional distribution runs all of the long trains in the same direction to create a uni-directional case. In cases where the number of long trains exceeds half the total traffic volume, the 100-0 directional distribution case exhibits strong directional preference by running as many long trains as possible in one direction, with a smaller number returning in the opposite direction as required to provide an even flow of railcars.

In the experiment design, each of these four factors has a specific number of values, or "levels" associated with it (Table 3.2). For example, the factor for the percentage of railcars in long trains was assigned four levels: 0, 25, 50, and 75 percent. The analyses performed in this experiment simulate factorial combinations of these different values.

Table 3.2: Experiment design factors and levels

Experiment Design Factors	Number of Levels	Level Specification
Percent Long Sidings	14	0, 4, 9, 13, 22, 30, 48, 52, 70, 78, 87, 91, 96, 100
Percent Railcars in Long Trains	4	0, 25, 50, 75
Directional Distribution	2	50-50 (bi-directional), 100-0 (uni-directional)
Freight Throughput	2	3,600 cars & 2,400 cars

Within the context of simulations, a key assumption is the strategy used when distributing the long sidings across the 240-mile route. An idealized approach is considered, where the number of long sidings corresponding to each factor level was distributed evenly along the route (Figure 3.1).
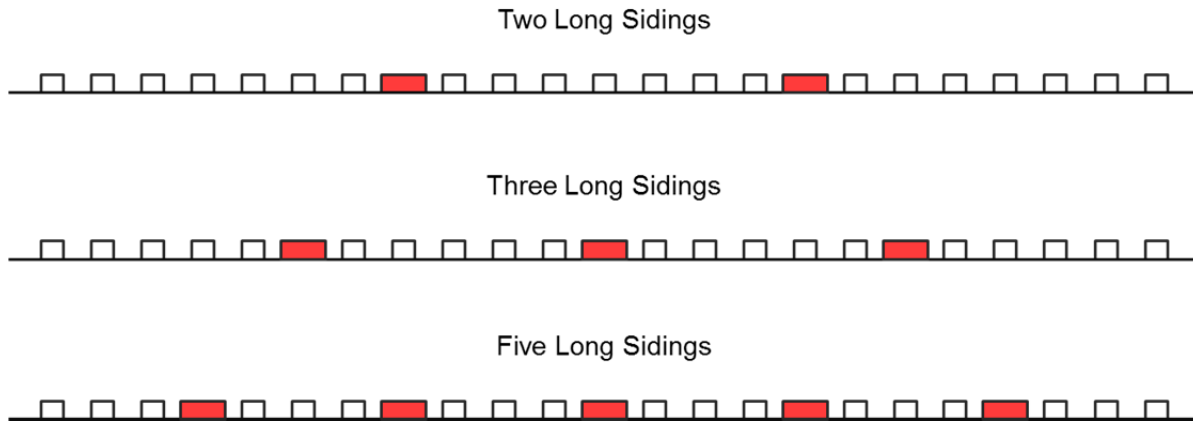
Figure 3.1: Balanced route distribution of long sidings for three example levels

A drawback of this distribution approach is that the pattern of long sidings does not represent, necessarily, a true progression of siding extensions that can be phased-in over time. For example, the locations of long sidings in the case of two and three long sidings cannot be built sequentially (Figure 3.1). A railroad cannot extend the first two long sidings and then later add a third and arrive at the same pattern of three evenly-distributed long sidings. They could, however, extend one long siding in the middle of the route initially, and eventually extend two

more (for a total of three) and still maintain a balanced route. Thus, to arrive at the evenly-spaced patterns of long sidings considered in this experiment, a railroad needs to first determine how many total long sidings they will require, and then build out accordingly.

To control for any difference in acceleration performance between short and long trains, the number of locomotives was proportionately increased for the long trains to maintain a constant horsepower-per-ton ratio. Thus the short 100-car trains operate with two locomotives while the long 150-car trains operate with three locomotives.

By combining the different factor levels from Table 3.2, 196 unique combinations were produced, each one corresponding to a simulation scenario in the experiment design matrix.

3.4 Results

After running simulations for the varying route and train characteristics described in the previous section, train delay data are exported from RTC and used to characterize the relationship between short sidings and long-train operation. For each individual simulation scenario in the experiment design matrix, train delay data is collected from five replications of five days of operations. Delay accumulated by individual trains during this 25-day period is averaged and normalized to produce a "delay per 100 train-miles" response for that element of the experiment design matrix. This average response for a simulation scenario is plotted as a single data point on the graphs that follow.

The results are divided into separate discussions for the simulations pertaining to bi-directional (50-50) long-train distribution and uni-directional (100-0) long-train distribution. The results for each operating pattern are eventually merged into a comprehensive discussion of their combined implications.

3.4.1 Bi-Directional Long-Train Operation

The results of simulating all cases with bi-directional long-train operations are illustrated for the 2,400-car throughput (Figure 3.2a) and the 3,600-car throughput (Figure 3.2b) by plotting delay per 100 train-miles on the vertical axis and percent long sidings on the horizontal axis.
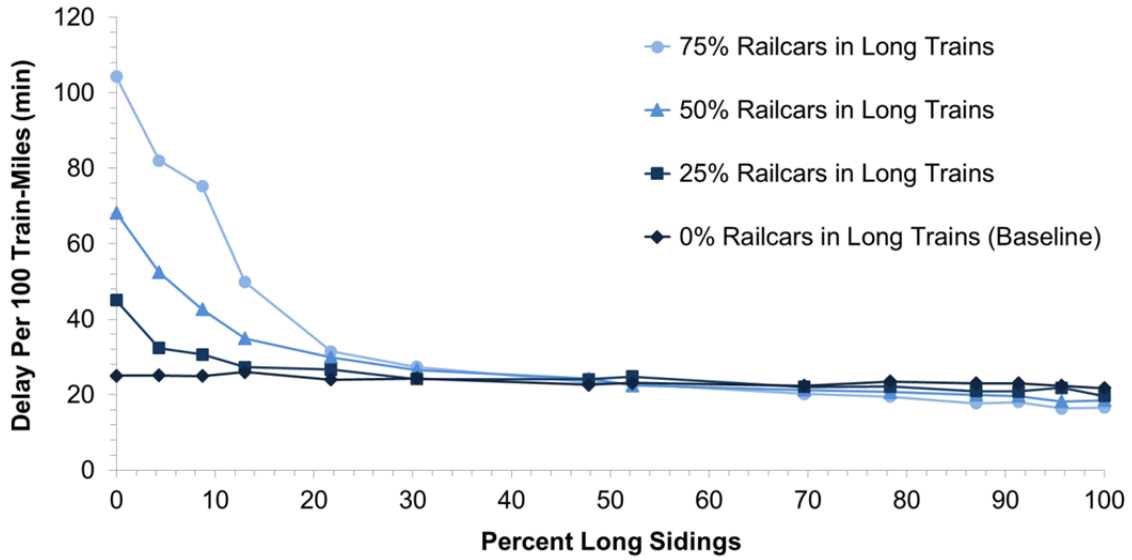
(a)
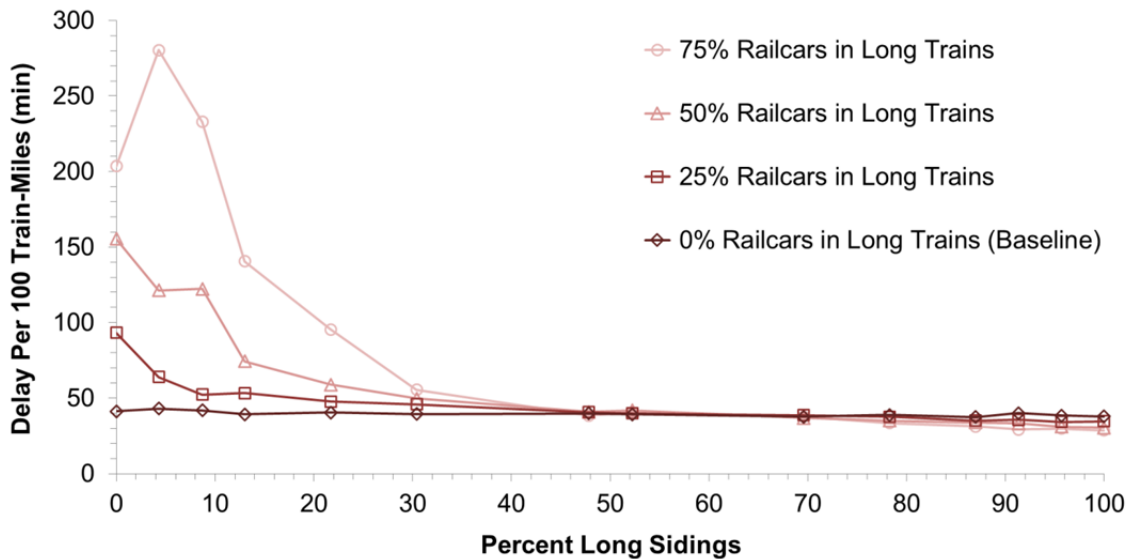
(b)

Figure 3.2: Delay as a function of percent long sidings and percent railcars in long trains, (a) 2,400-car throughput and (b) 3,600-car throughput

The different curves represent the different percentages of railcars in long trains. The delay curves in both figures exhibit three zones of behavior: one at low percent long sidings (between 0 and 40 percent), one near the middle range of percent long sidings (40 to 70 percent), and one at high percent long sidings (between 70 and 100 percent).

Beginning on routes with a low percentage of long sidings, a relatively steady downward trend in delay, almost exponential in nature, dominates the space. The exception is the curve for routes containing 50 or 75 percent railcars in long trains, where larger variation in train delay within individual simulations leads to a slight fluctuation in the data on the left side of both figures. As the number of long trains being operated on the route increases, the train delay for a given level of percent long sidings increases.

In situations where there is a high percentage of long trains and a low percentage of long sidings, large train delays are observed. On routes with no long sidings, converting 25 percent of the traffic to long trains can double average train delay while converting 75 percent of traffic can increase average train delay by a factor of four. The lack of locations where two long trains can meet creates a dispatching phenomenon where long trains are fleeted across the entire route successively. This form of fleeting leads to excessive delay as long trains are held in terminals until several long trains are ready to depart in rapid succession. These fleets also disrupt short-train movements on the line, as the short trains stop for longer periods of time in passing sidings to meet multiple long trains. The result is an inconsistency in the operating pattern that in turn causes high variability in delay data. This fluctuation, however, is short-lived and delay reduction declines to a single, critical point as more long sidings are added.

For both throughput volumes, the delay trends converge when slightly less than 50 percent of the passing sidings are extended to long sidings. At this point, delay for the 25, 50,
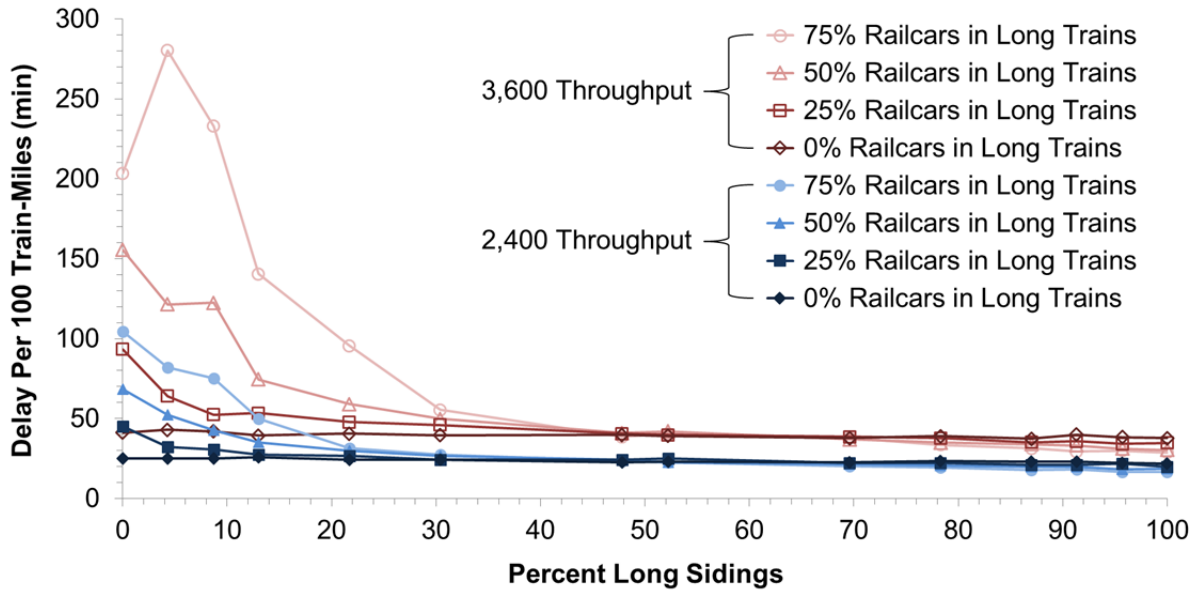
and 75 percent railcars in long train scenarios is equal to the 0 percent railcars in long trains baseline scenario. In the baseline scenario, all trains are short and can therefore use any passing siding for a meet. The baseline scenario is thus expected to show little delay response to the addition of long sidings, since short trains are essentially indifferent to siding length.

The point where the trends associated with long-train scenarios converge with the baseline level of service may be the most critical piece of information to planners and engineers in charge of siding extension and construction programs. The implication is that routes with roughly half of their sidings extended to handle long trains will avoid any delay-based consequences of operating long trains on the route. Restated, to operate with a high percentage of long trains, only half of the sidings on a route need to be extended in order to maintain the baseline level of service. These results are solely based on the tested combinations of train lengths and balanced siding distributions presented earlier. The results might change if different combinations of train lengths or build-out patterns were employed, as will be investigated in the next chapter.

Past this critical point, the data between roughly 70 percent and 100 percent long sidings also exhibit an interesting trend. As larger numbers of passing sidings on the route are extended to support meets between two long trains, delay becomes almost entirely linear, and there are little or no negative effects of long train operation on route delay. Delay for the case of 75 percent railcars in long trains is actually the lowest, while delay for the base case with all short trains is the highest. Over this range, the more long trains operating on the line, the lower the simulated delays along the route. As more long trains are operated with total throughput held constant, the total number of trains on the route decreases. With a smaller train count, there is an expectation of reduced delay, as observed in the right-hand tails of Figures 3.2a and 3.2b. In this

range, the capacity efficiencies afforded by the operation of fewer, longer freight trains are fully realized. For the sake of comparison, the delay curves for the two throughput levels are superimposed to emphasize the consistency in delay patterns for the two different freight throughput volumes and corresponding combinations of short and long trains (Figure 3.3).

Figure 3.3: Delay as a function of percent long sidings and percent railcars in long trains, overlaid 3,600-car and 2,400-car throughputs

Observation of the variance in delay response from the simulation scenarios (Figure 3.4) indicates a large variance for delay values obtained at a percentage of long sidings less than 20 percent. This corresponds to the region of delay data variability on the left side of Figure 3.3, as was discussed previously. While the shape of each delay variance curve in this region is erratic and somewhat inconsistent, the observation that all variances converge to small values at roughly 20 percent long sidings means small differences in delay values at higher levels of percent long sidings are more significant, as the delay variance is small across this region.
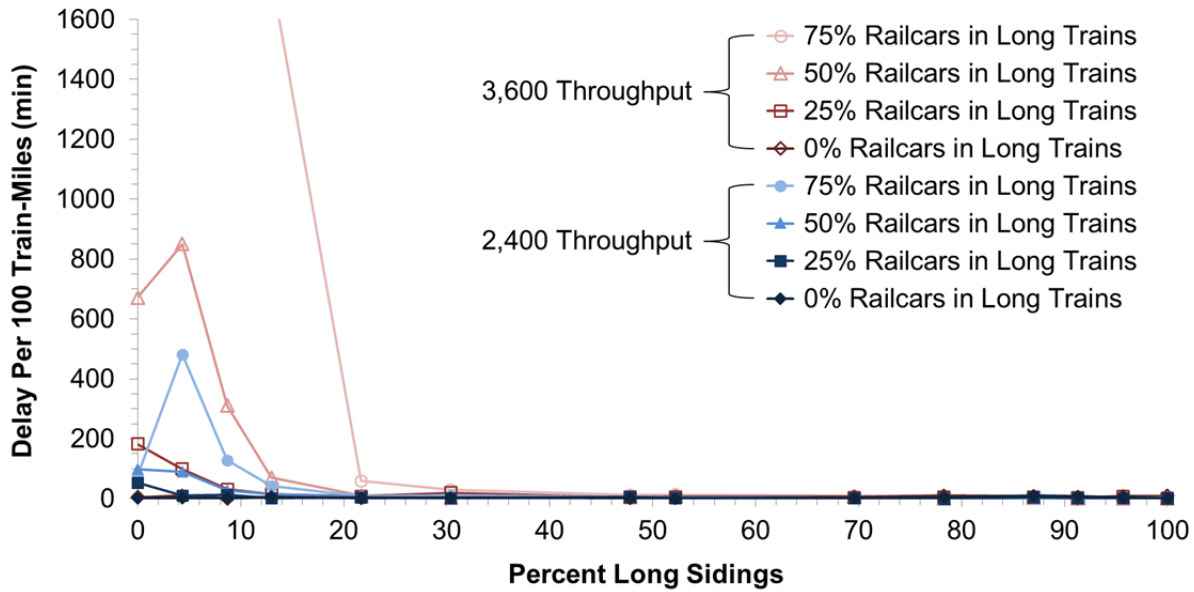
Figure 3.4: Delay variance as a function of percent long sidings and percent railcars in long trains, 3,600-car and 2,400-car throughputs

3.4.2 Uni-Directional Long-Train Operation

The previous section suggests that bi-directional operation of long trains can be supported with minimal delay impact if 50 percent of the sidings on a route are extended, based on the assumption that long sidings are spaced out evenly along the route. Although half of the existing passing sidings do not need to be altered, building the required siding extensions still represents a sizeable capital investment. Also, there may be environmental, engineering, or construction constraints that prevent the extension of certain passing sidings, potentially disrupting the even distribution of long sidings. Although the effect of an uneven long-siding distribution is the subject of future study, such a scenario could cause additional delay, requiring extra siding extensions to match the original all-short-train base case.

To avoid this investment, railroads may elect to operate long trains in a single direction to avoid any meets between two long trains (Dick & Clayton 2001). Since a long train on the main track can pass a short train on an existing siding, operating long trains in only one direction does

22

not require any additional passing siding infrastructure. However, operation of long trains in a single direction does introduce complications. Since the number of trains operating in each direction is unequal, it creates an asymmetry in crew requirements. This introduces the expense of extended layovers or deadheading crews back to the origin terminal to match the uneven train flow. Similarly, on certain routes, the required number of locomotives for the short and long trains may be such that there is an imbalance in locomotive demand in each direction. This will reduce locomotive utilization and increase the number of non-revenue locomotive deadhead miles required to reposition equipment. Running a long train in one direction without a corresponding long train in the other direction can also complicate train planning and block-to-train assignment. It also means that unit and shuttle trains, which benefit the most from the efficiency of long trains, cannot operate as long trains for both legs of their round-trip journey. Instead, the long unit trains must be broken up and recombined at either end of the trip. Finally, uni-directional operation of long trains dictates that at most, only 50 percent of traffic (i.e. all of the traffic moving in one direction) can move in long trains. Conversion of additional traffic to long trains will require that some returning trains also be operated as long trains.

Despite these complications, there are economic and productivity benefits to implementing uni-directional long-train operation. The simulation results for uni-directional cases indicate that operating less than 50 percent of total traffic as long trains in a single direction has no impact on train delay (Figure 3.5). Since the long trains all travel in one direction, there are no two-long-train meets and the delay for these cases matches that of the 100 percent short-train base case (or shows a slight improvement due to reduced total train count).

However, if additional long trains are run in the return direction, the results echo those obtained for the bi-directional scenario presented earlier, as shown by the 75 percent railcars in

long trains data series in Figure 3.5. This condition is no longer a fully uni-directional case, but rather one with a directional preference. When the delay curve for directional preference at 75 percent railcars in long trains is superimposed on the equivalent bi-directional results shown previously in Figure 3.2b, the curve shows the familiar three-staged delay behavior (Figure 3.6).
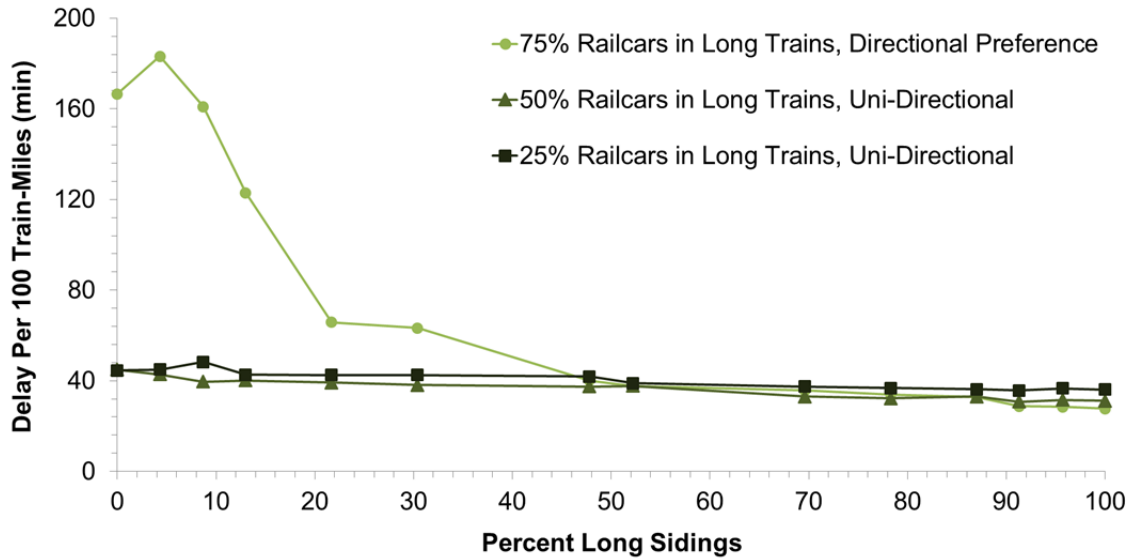
Figure 3.5: Delay as a function of percent long sidings and percent railcars in long trains; Uni-directional operation compared against directional preference, 3,600-car throughput
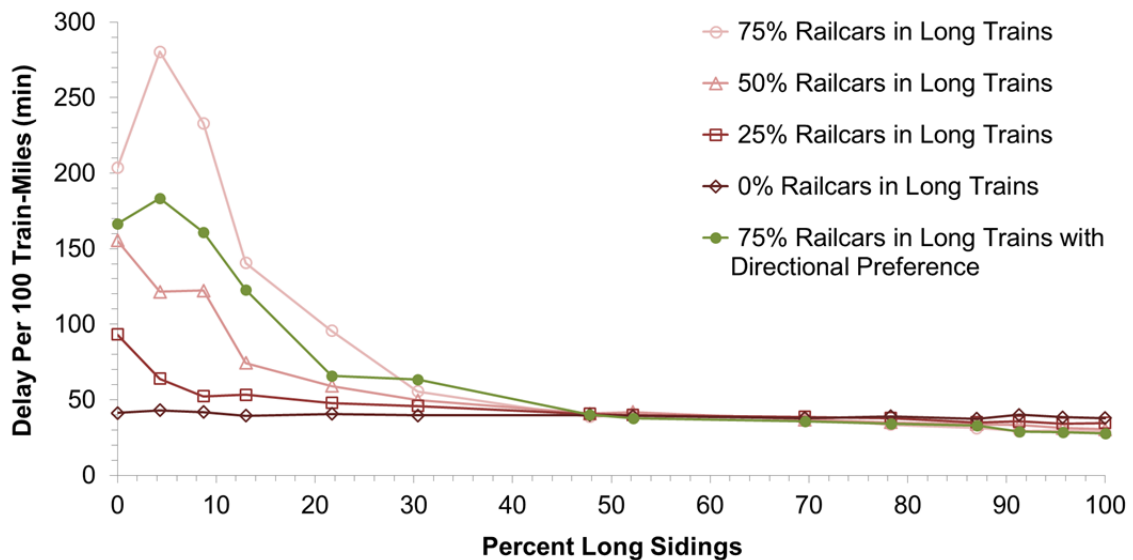
Figure 3.6: Delay as a function of percent long sidings and percent railcars in long trains; Directional preference compared against bi-directional operation, 3,600-car throughput

24

This is not unexpected since this case involves the operation of long trains in both directions, but with an imbalance (six in one direction, twelve in the other). Most notably, however, this new delay curve converges to roughly the same 50 percent long siding mark observed in the true bi-directional data. This result suggests the number of long sidings required to mitigate delay from long-train operation is independent of the exact directional distribution of long trains.

At low percent long sidings, the delay curve for the case of long train directional preference lies below that of its bidirectional equivalent (Figure 3.6). This result is intuitive since the bi-directional case, with nine long trains operating in each direction, has the potential for a maximum of 81 (9×9) two-long-train conflicts. The case with directional preference, with six long trains in one direction and twelve in the other, only has the potential for a maximum of 72 (6×12) conflicts. Thus, to minimize the impact on train delay, consideration should be given to running a majority of long freight trains in one direction until sufficient numbers of long sidings can be constructed to operate equal numbers of long trains in each direction.

3.5 Conclusions and Future Work

North American railway operations have experienced a dramatic shift with the advent of distributed power, spurring increased use of longer freight trains to transport cargo along existing rail corridors. The economical and operational efficiencies that longer freight trains provide are constrained by the length of many passing sidings. This study presents a simulation approach to evaluate operations on a representative single-track line under various combinations of freight throughput, percent railcars in long trains, percent long sidings, and the directional distribution of long trains operating on the route.

Results indicate that routes with roughly 50 percent long sidings exhibit no delay-based consequences of running long trains. This suggests that to operate with a high percentage of long

trains, only half of the sidings on a route need to be extended in order to maintain the baseline level of service (i.e. the average train delay with no long trains in operation). On routes with more than 50 percent long sidings, total train count takes precedence over the ratio of long to short trains in determining train delay. Results also indicate a similarity in delay-reduction patterns regardless of whether long trains operate with a 50-50 directional distribution or with directional preference. This finding also highlights the improved delay characteristics associated with running a majority of long trains in one direction, as opposed to 50-50 bi-directional operations. When running fewer long freight trains, uni-directional operation has no adverse effects on train delay while simultaneously minimizing infrastructure investment. These findings can serve as general guidelines for developing siding extension and construction programs while simultaneously facilitating the efficient operation of long freight trains.

The next chapter investigates a broader range of ratios of long to short train lengths to determine if the free-flow point of 50 percent long sidings varies, or is a fundamental property of single-track lines.

The routes considered in this chapter are idealized and, as with other delay and capacity relationships, the trends between percent railcars in long trains and percent long sidings may not hold for routes with uneven siding spacing (Atanassov et al. 2014) – a condition explored in Chapters 5 and 6. This research also only considered routes with homogeneous freight traffic. Since it has been shown previously that introducing traffic heterogeneity can alter capacity relationships (Dingler et al. 2012), introducing heterogeneity to the simulations in the form of passenger trains may alter the results. Future simulation experiments will include these and other factors to investigate these possibilities.

Finally, the analysis in this study assumes that all yards, terminals, and loading/unloading facilities have the capacity to handle long 150-car trains. Just like passing sidings, yard and terminal tracks have been constructed to match the shorter trains of previous eras, and balloon loops at bulk freight transload facilities are designed for a particular design train length (Dick & Brown 2014, Dick & Dirnberger 2014). Without adequate infrastructure, long trains may affect the delay and capacity of these facilities with an overall negative impact on network performance that offsets gains from reduced train counts. These terminal effects may also be worthy of future investigation.

CHAPTER 4
LONG-TRAIN REPLACEMENT RATIO AND
INFRASTRUCTURE INVESTMENT

4.1 Introduction

Use of longer train consists is hindered by existing track infrastructure in North America, where mainlines are predominantly single track with passing sidings. Many passing sidings lack sufficient length to hold trains in excess of 100 railcars, effectively setting an upper bound on North American freight train lengths. Typically, existing passing sidings range from 6,000 to 7,500 feet, a length sufficient to hold about 100 to 125 railcars. By contrast, most new siding construction projects range from 9,000 to 10,000 feet, enabling operation of 150-car trains with seven locomotives in a distributed-power configuration.

The research presented in this chapter aims to build upon the relationships generalized in Chapter 3 through a more comprehensive analysis of the infrastructure required for routes operating with different combinations of short and long train lengths, as expressed by the "train replacement ratio". The results of this study can be used to develop a better understanding of the interaction between train delay, the lengths of passing sidings, and the relative lengths of trains operating on a particular freight corridor. Ultimately this knowledge can help streamline the decision-making process associated with the implementation of long-train operations and rail infrastructure expansion programs to extend passing sidings or construct new longer sidings.

4.2 Notation

The concept of a "train replacement ratio" is frequently referred to in the following sections. For the purpose of this research, train replacement ratio is defined as the ratio of the length of long trains on a route as compared to short trains on the same route. For example, if a route operates long trains of 150 railcars and short trains of 100 railcars, two of the long trains can move the same amount of freight (railcars) as three short trains. In other words, two long trains can replace three short trains and contribute to a reduced total train count. The corresponding train replacement ratio is 3:2. In general, the larger the replacement ratio, the larger the disparity between long and short train sizes. Small increases in train length correspond to small train replacement ratios.

4.3 Methodology

The overarching simulation methodology used throughout this research builds on the work presented in Chapter 3, and is based on a representative single-track, heavy-haul route (Table 3.1). The study presented in this chapter simulates additional freight-train lengths in different combinations to achieve various replacement ratios with the aim of understanding how the number of long sidings required to maintain a certain level of service (train delay) on a route is related to the train replacement ratio. The new train lengths under consideration are detailed later in this section.

In the experiment design, the number of locomotives assigned to each train is varied in proportion to its length to maintain a constant horsepower-per-ton ratio. This proportional addition of power to the longer trains helps control for subtle differences in acceleration and braking performance that might cause additional congestion and delay, thereby confounding comparisons of the simulation results. Two siding lengths are specified: a shorter length to

represent current passing siding conditions and a second longer length to represent passing sidings that have been extended.

The research question this chapter seeks to answer is: does changing the ratio of long and short train lengths have an effect on the amount of siding investment required to restore baseline levels of service (a condition defined by exclusive short-train operations) or is the number of required sidings independent of this factor? The experimental design for the Rail Traffic Controller (RTC) simulations is comprised of three main variable factors: Percent Long Sidings, Percent Railcars in Long Trains, and Train Replacement Ratio. The first two factors are defined in section 3.3, and replacement ratio was defined earlier in section 4.2.

Percent long sidings is the fraction of total sidings on the route that have been extended from the base length of 100 railcars to be longer than the length of the long trains, in this case either 120 or 150 railcars. A key assumption is that an idealized strategy was used in distributing long sidings along the simulated route. Long sidings were always distributed evenly such that the route remained balanced from an infrastructure perspective. Percent railcars in long trains is the fraction of total railcars on the route moving in long trains.

To achieve a range of replacement ratios, different combinations of 150-, 120-, 100-, 75- and 50-car trains were used. The 6:5 and 3:2 ratios (120 & 100-cars trains, and 150 & 100-car trains, respectively) represent common operational situations facing North American heavy-haul operators as they increase the length of unit trains. The 2:1 and 3:1 ratios use artificially short train lengths that are not truly representative of current operating conditions but are used to extend the trends and relationships apparent in the results without resorting to simulating extremely long 300-car trains.

Each of these three factors has a specific number of values, or "levels", associated with it (Table 4.1). For example, the factor Percent Railcars in Long Trains was subdivided into four levels: 0, 25, 50, and 75 percent railcars in long trains. The analyses performed in this study are based upon simulated factorial combinations of these different values.

Table 4.1: Experiment design factors and levels

Experiment Design Factors	Number of Levels	Level Specification
Percent Long Sidings	14	0, 4, 9, 13, 22, 30, 48, 52, 70, 78, 87, 91, 96, 100
Percent Railcars in Long Trains	4	0, 25, 50, 75
Train Replacement Ratio	4	6:5 (100-car short & 120-car long) 3:2 (100-car short & 150-car long) 2:1 (75-car short & 150-car long) 3:1 (50-car short & 150-car long)

The throughput volume considered in this study remained fixed at 2,400 railcars per day, and the directional distribution of all traffic along the route was 50-50, i.e. evenly distributed in both directions. There are particular efficiencies afforded by uneven directional running of long trains (such as reduced meets between long trains) that were previously discussed in Chapter 3. However, the goal of this study is to focus on the case where unit-train cycles with fixed train consists make it impractical to only run long trains in a single direction.

This study does not consider the ability of unit-train loading facilities, unloading facilities, and any intermediate staging and inspection yards to support the operation of longer trains. While the focus is on mainline single-track operations, in practice, additional terminal infrastructure investments may be required to establish tracks and loops of sufficient length to support long-train operation.

4.4 Results

Each scenario in the experiment design was simulated with RTC to generate train delay output. For each individual simulation scenario in the experiment design matrix, train delay data is collected from five replications of five days of operations. Delay accumulated by individual trains during this 25-day period is averaged and normalized to produce a "delay per 100 train-miles" response for that element of the experiment design matrix. This average response for a simulation scenario is plotted as a single data point on the graphs that follow. Simulation results for the combination of scenarios presented in the methodology section were compiled to highlight the relationship between track infrastructure (number of siding extensions) and the number (percent railcars in long trains) and relative length (replacement ratio) of long trains.

The results support the generalized relationship between percent long sidings and train delay introduced in Chapter 3. Two types of behavior are exhibited when long trains are operated on single-track lines with short sidings (Figure 4.1).
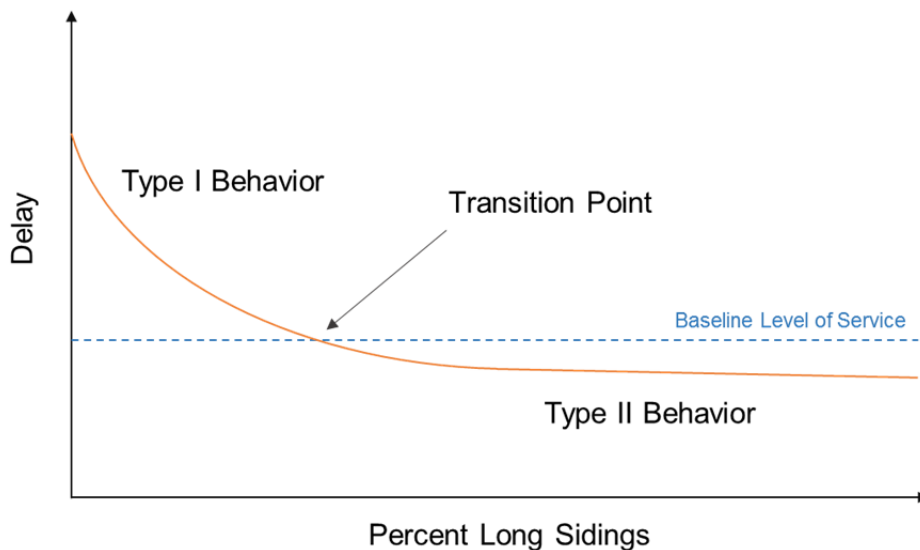

Figure 4.1: Delay behavior for long train operations

Type I behavior describes the condition where the extra delay associated with the inflexibility of long-train meets on routes with inadequate numbers of long sidings outweighs the reduced train delay resulting from the reduction in train count afforded by the long trains. The net result is that the route operates with a higher average train delay than the baseline condition of pure short-train operations, even though the baseline has a higher total train count.

Type II behavior describes a condition where there are enough long sidings providing flexibility for long-train meets that the benefits of reduced train count are realized. Under these conditions, the route operates with lower average delay than the baseline condition even though the train dispatcher is still constrained by the number of sidings usable for meets between long trains. Although some long trains may still be delayed, most will exhibit acceptable amounts of train delay due to the reduced number of trains on the line.

The "transition point" between these two types of behavior indicates the level of siding extension investment required to mitigate the delay increases resulting from long-train operations, and to return the route to its baseline level of service. The experiment matrix is designed to determine if the number of siding extensions at the transition point is related to the train replacement ratio, as will be explored quantitatively in subsequent paragraphs.

To quantitatively illustrate the delay behavior in Figure 4.1, the simulation results for a 3:2 replacement ratio (150-car long trains & 100-car short trains) at 2,400 cars per day highlight the relationship between route capacity (in the form of train delay) and percent long sidings as a function of percent railcars in long trains (Figure 4.2a). In all cases, the "0% Railcars in Long Trains" series represents the baseline case of all-short-train operations. Since two short trains can already meet at any siding along the route, they are not sensitive to creation of long sidings. Thus this series shows the expected constant response at the baseline level of service. The baseline
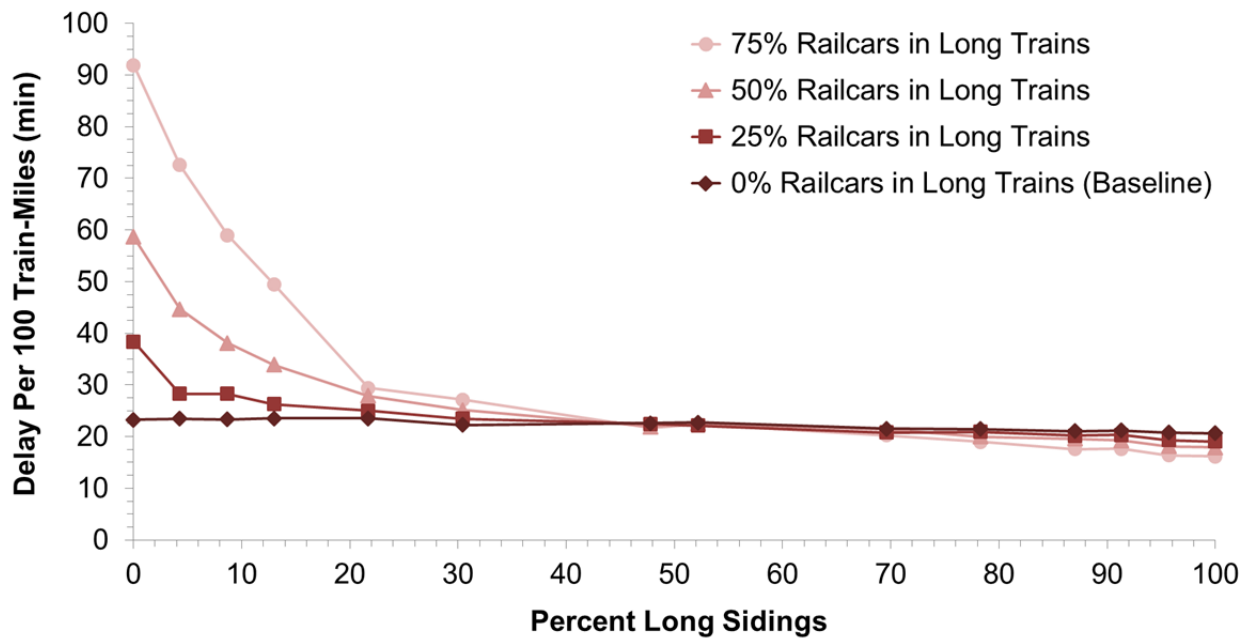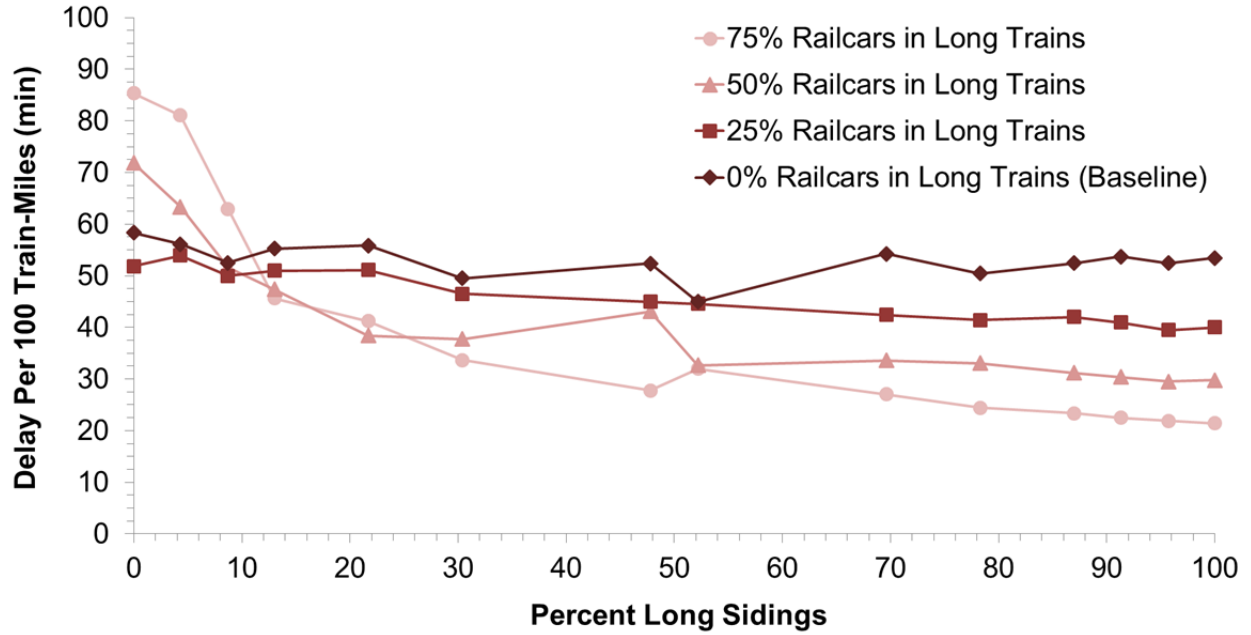
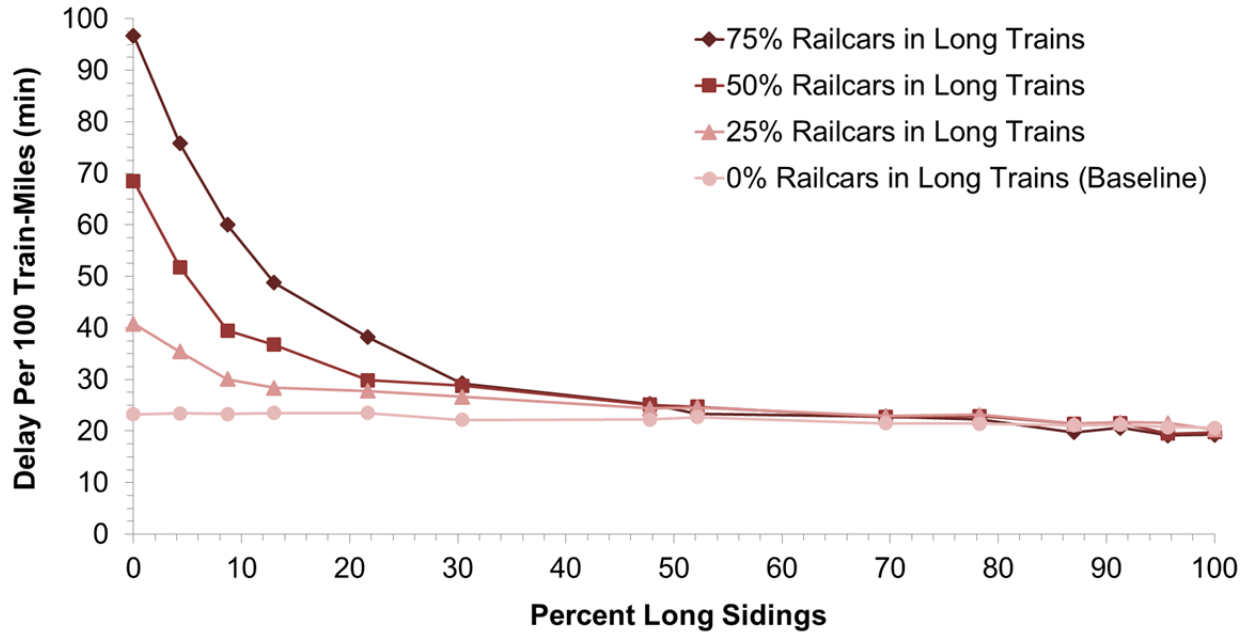**(a)**
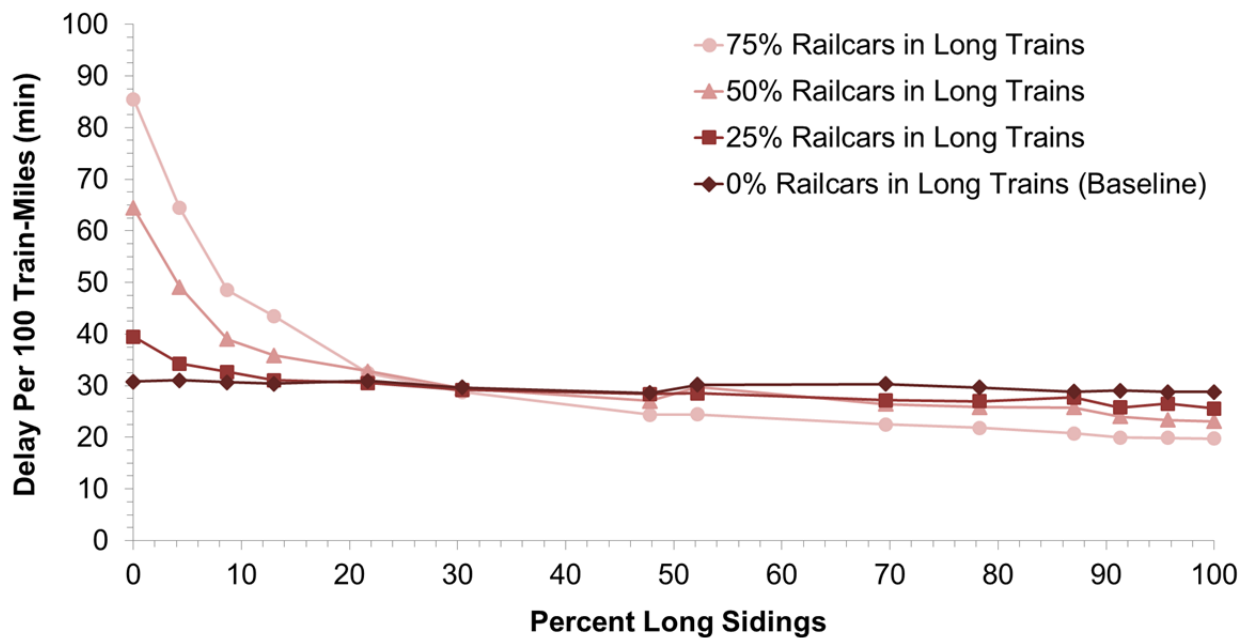
(b)

Figure 4.2: Delay characteristics based on a replacement ratio of (a) 3:2, (b) 2:1

(c)

(d)

Figure 4.2 (cont.): Delay characteristics based on a replacement ratio of (c) 3:1, (d) 6:5

level of service is, however, a function of the initial number of short trains (as determined by the total throughput and the length of the short trains in each experiment scenario).

The form of Figure 4.2a falls into three "zones" of delay response – one between 0 and 50 percent long sidings, one near 50 percent long sidings, and one between 50 and 100 percent long sidings. The first zone exhibits Type I behavior where, as described for Figure 4.1, a rapid decrease in delay is evident as long sidings continue to be added along the route. Also, curves corresponding to cases with relatively high percent railcars in long trains show the highest average train delay. The third zone, on the other hand, shows Type II behavior in the form of lower delay values for cases that include higher percent railcars in long trains and, therefore, a lower overall train count.

In Figure 4.2a the point of convergence of all lines near 50 percent long sidings (the "transition point") indicates that, for this combination of train lengths and traffic volume, in order to operate with a high percentage of long trains, only half the sidings on a route need to be extended to maintain the baseline level of service (defined by existing short-train operations). At levels of percent long sidings above this transition point, the economies of scale of long-train operations result in reduced delay for cases with more long trains and a correspondingly lower train count. The broader range of replacement ratio values present in the experiment matrix was designed to test the consistency of this transition point.

The results from simulating the same volume of 2,400 railcars per day but with a replacement ratio of 2:1, 3:1, and 6:5 (Figures 4.2b, 4.2c, and 4.2d) exhibit the unique behavior of the transition point across different replacement ratios. It is apparent that the transition point, originally near the 50 percent long-siding mark in Figure 4.2a, has shifted to the left in Figure

4.2b and is nearer to the 30 percent long-siding mark. This suggests that the transition point varies with the ratio of train lengths being operated along any one particular route.

That the transition point moved to the left from Figure 4.2a to Figure 4.2b implies that the larger the difference between long and short train lengths (i.e. replacement ratio), the less siding investment is required to reach the transition point between Type I and Type II behavior. From a practical standpoint, the results in Figure 4.2b imply that, in order to achieve economies of scale from running longer trains at a 2:1 replacement ratio, roughly 30 percent of sidings need to be extended on a route in order to accommodate longer trains without additional delay.

Again comparing Figure 4.2b to Figure 4.2a, the lines exhibiting Type II behavior to the right of the transition point in Figure 4.2b are spaced farther apart, indicating a greater delay reduction resulting from the operation of long freight trains in instances where long sidings are more frequent. This result confirms the expectation that operation of longer and longer freight trains compared to existing short trains will offer greater improvements in operational efficiency based on greatly-reduced train count alone.

This effect of replacement ratio on the transition point is further supported by the results of the scenarios with the other two replacement ratios (Figures 4.2c and 4.2d). The transition point moves furthest to the left for the highest replacement ratio (3:1) and furthest to the right for the lowest replacement ratio (6:5).

To determine if there was a specific form to the observed trend between train replacement ratio and the corresponding level of siding investment implied by each transition point, the two were plotted (Figure 4.3). The "level of siding investment" here refers to percent long sidings.
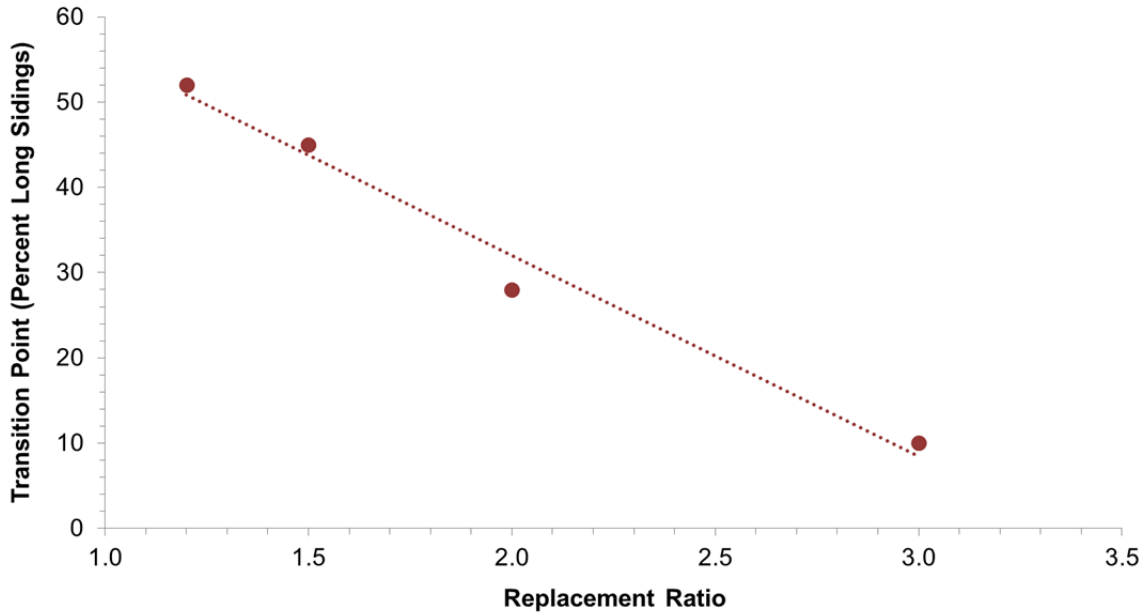
Figure 4.3: Transition point as a function of train replacement ratio

The resulting plot of infrastructure investment at the transition point as a function of train replacement ratio (Figure 4.3) shows a fairly linear relationship, with the amount of sidings at the transition point decreasing as replacement ratio increases. These results can be applied to railway industry practice in that they provide some insight, at least from a delay perspective, into siding extension programs or, alternatively, relative train-length optimization.

For example, consider the case of a railroad that wants to expand long-train operations but only has enough capital to extend a certain percentage of their passing sidings. They can use the relationship derived here and the maximum percent of long sidings dictated by budget constraints to get a better sense of how much their trains would need to be lengthened in order to maintain their current level of service. Alternatively, if a target train length (and corresponding replacement ratio) has already been proposed, a more streamlined estimation of the required number of siding extensions can be developed as part of a capital plan. Since the relationship

appears to be linear, there does not appear to be an optimal point of diminishing returns in this regard.

The idea of linearity at extreme points in Figure 4.3 can, however, be argued against conceptually. For example, if the replacement ratio was 1.01 (101-car long trains, 100-car short trains), it can be expected that since there is little reduction in total train count, almost all of the sidings along a route would need to be extended in order to maintain the baseline level of service. Alternatively, for a hypothetical replacement ratio of 12 (1,200-car long trains, 100-car short trains), it might be such that no sidings need to be lengthened since only two trains need to be run to achieve a 2,400-car throughput. These conceptual data points would not follow the linear relationship suggested in Figure 4.3. Linearity as shown in Figure 4.3 may therefore just be a function of the limited range of values tested, with extremities potentially highlighting a more complex relationship. In either case, observation of more scenarios that test a broader range of replacement ratios can test the validity of the relationship observed thus far. However, the current result covers the most practical long-train replacement scenarios that are being considered in practice and can already help streamline siding-extension and train-lengthening programs in the railway industry.

4.5 Conclusions

The economic and operational merits afforded by long train operation are often constrained by inadequate passing siding lengths in North America. This problem necessitates the need for infrastructure expansion in the form of either siding extension programs or construction of additional longer passing sidings. The research presented in this chapter uses a simulation approach to analyze the relationship between train replacement ratio (i.e. the ratio of the length of long trains on a route relative to short trains) and required siding investments.

Results show a declining linear relationship between train replacement ratio and the point where siding investments on a route mitigate additional train delays introduced by operating long freight trains that exceed the length of passing sidings. The larger the replacement ratio, the fewer passing siding projects that must be completed to achieve the economies of scale expected from long-train operations. These findings can streamline the planning process by providing insight into the scope and magnitude of siding extension programs required in anticipation of longer freight-train operations and their desired return on investment.

CHAPTER 5
SINGLE TO DOUBLE TRACK: INFLUENCE OF SIDING SPACING

An earlier version of this research appears in:

Atanassov, I., C.T. Dick & C.P.L. Barkan. 2014. Siding spacing and the incremental capacity in the transition from single to double track. In: Proceedings of the 2014 Joint Rail Conference, Colorado Springs, USA.

5.1 Introduction

The majority of the railway network in the United States is single-track mainlines with passing sidings. Single-track lines impose constraints on handling conflicting train movements, causing service levels to decline as traffic volumes grow. With demands for freight and passenger services forecasted to increase, it will be necessary to expand rail infrastructure to accommodate the additional traffic. While methods for increasing capacity on rail corridors vary, common approaches involve extension of existing sidings (e.g. a "super siding"), or construction of new sidings (as discussed in the previous two chapters). While these steps provide initial solutions to the problem, if traffic continues to grow it may become necessary to install sections of double track to accommodate increasing volume while providing an adequate level of service.

The analysis that follows aims to characterize incremental delay-reduction trends resulting from double-tracking corridors with different distances between passing sidings. Delay in this study serves as a simultaneous measure of capacity and level of service. Previous research has considered this question on idealized lines with evenly-spaced passing sidings. The objective of this study is to determine if the same trends are exhibited by a siding connection strategy for a more realistic scenario with a mixture of siding spacings along a corridor. This is much more typical of real-world physical and engineering constraints. While there are many factors to consider in planning for additional infrastructure, identifying trends in double-track build-out strategy is meant to improve the planning process by helping to generally identify the types of

projects with the greatest potential benefits. These strategies should be considered along with other delay-causing factors such as local switching work, yard locations, and grades that may make double track more attractive on some mainline segments than others. With a smaller number of prioritized project alternatives, railroads can better utilize their modeling, planning, and engineering resources to conduct more detailed analyses to make a final selection between the remaining track expansion options.

5.2 Background

The delay characteristics for single-track mainlines have been well-covered in existing literature, and research has been extended into studies on the delay benefits of partial double-track installation (Lindfeldt 2006, Lindfeldt 2012, Sogin et al. 2013a). The subsequent analyses provided in this chapter are an extension of results obtained by Sogin et al. (2013a), who found that for idealized corridors with even, 10-mile siding spacing, there was a linear reduction in train delay as a function of percent double track for several freight traffic volumes. The reductions in delay resulting from double-track installation are consistent with the idea that train meets are the primary cause of delay on single track and that double track allows a larger proportion of trains to avoid meets (Dingler et al. 2010).

Sogin et al. (2013a) used Rail Traffic Controller (RTC) to conduct simulation experiments that investigate different strategies for transitioning from single to double track. They found that the optimal strategy was an alternating build-out approach (Figure 5.1) that will be used in the analysis that follows. The alternating strategy involves setting four to six points along the length of a route from which a second mainline track would be built out in both directions progressively while connecting existing passing sidings.

Figure 5.1: Alternating build-out strategy where four midpoints are chosen along the corridor, and second mainline track is built out in each direction

A part of the analysis presented in the following sections supplements Sogin et al.'s (2013a) results for 10-mile siding spacing. Experiments were conducted for a route with longer, 16-mile siding spacing. Conclusions drawn from this analysis are then extended further by application to more realistic scenarios where initial siding spacing is non-uniform.

5.3 Methodology

Different methodologies were employed for the two studies presented in this chapter: one for delay characteristics of routes with 16-mile siding spacing, and one for the optimal double-track installation strategy for a route with non-uniform siding spacing. These methodologies are detailed separately in the following sections. In practice, it is the running time between sidings and not the siding spacing distance that controls the capacity of a single-track line. However, both the grade and maximum track speed on all sections of the hypothetical line are uniform, resulting in the same operating speed. Thus, in this analysis, the distance between passing-siding centers can be used as a proxy for running time between sidings.

5.3.1 Impact of Initial Siding Spacing

In order to identify how increasing the initial distance between evenly-spaced sidings affects the benefits of double-track installation, two models were specified and simulated in RTC to generate comparative delay characteristics (Table 5.1).

Table 5.1: Model parameters for two routes with differing initial siding spacing

Route & Train Characteristics	Model 1	Model 2
Siding Spacing	10 miles	16 miles
Minimum Percent Double Track	~ 19%	~ 12%
Maximum Percent Double Track	100%	100%
Trains Per Day Range	8-60 (8 Levels)	8-64 (10 Levels)
Traffic Composition	100% Freight	100% Freight
Locomotives	SD70 (x3)	SD70 (x3)
Number of Cars	115 Hoppers	115 Hoppers
Length	6,325 feet	6,325 feet
Mass	16,445 tons	16,445 tons

Both models are based on an identical, 240-mile route subject to homogeneous freight train operations; the models differ only in the initial arrangement of their sidings (i.e. the number and spacing between them). The sidings in each model were then incrementally connected using the alternating strategy (Figure 5.1) until the entire 240-mile route was composed of a two-track mainline with universal crossovers at one end of each former siding location.

Model 1 was simulated by Sogin et al. (2013a), who found a linear reduction in train delay as a function of percent double track. Model 2 is constructed and simulated in this work as a means of identifying the difference in delay patterns for a route that has longer bottleneck sections (i.e. single-track sections), due to sidings being spaced farther apart initially. The term "Levels" in the Trains Per Day row in Table 5.1 is used to differentiate the exact number of train volumes considered; for example, ten levels between 8 and 64 trains per day indicates that there were ten distinct train volumes modeled within that range.

Siding spacing and the number of sidings are the only differences in the two models: the difference in "Minimum Percent Double Track" is a mathematical result of the fixed route length and the increased siding spacing resulting in fewer initial sidings. The jump to ten levels of traffic volume in Model 2 as opposed to eight levels in Model 1 is used to improve the detail in

the results. The characteristics of the simulation process and calculation of average train delay for each model are described in Chapter 2.
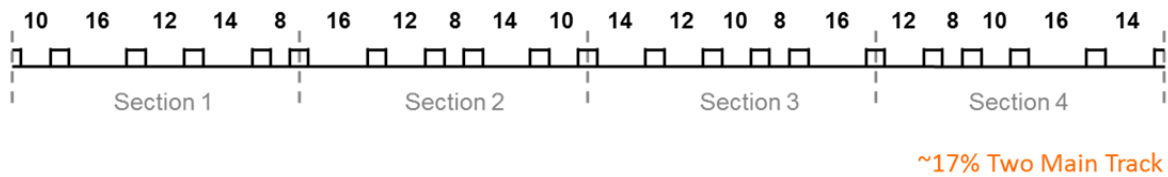
5.3.2 Variable Siding Spacing and Connection Strategy

The experiment on siding arrangement described in the previous section is designed to provide a better understanding of the delay characteristics of routes with a sparse arrangement of sidings compared to evenly-spaced sidings at closer intervals. However, single-track routes will rarely, if ever, have such ideal, evenly-spaced sidings due to a variety of engineering, operational, environmental, geographic, land use, and historical constraints. In order to investigate if the same linear delay-reduction trends identified by Sogin et al. (2013a) hold for a double-track installation strategy on corridors with a more realistic, non-uniform siding arrangement, a new set of model parameters were created.

In general, there are numerous strategies that can be employed when selecting the order to connect existing sidings to create double-track sections. The most intuitive strategy, taking local variation in construction cost out of consideration, would be to connect adjacent sidings that are the farthest apart first. Such a strategy ensures that the longest bottleneck sections are removed from the route soonest, presumably leading to the greatest reduction in delay. The goal of the following experiment is to determine what sort of siding connection strategy most effectively reduces train delay. In order to provide the greatest potential contrast in delay response, the two build-out strategies that were tested are the *short-to-long* strategy, where the sidings spaced closest together are connected first, and the *long-to-short* strategy mentioned above, where the sidings spaced farther apart are connected first (Figure 5.2). These two build-out strategies are implemented on the initial route layout with squares representing existing

passing sidings and the numbers above the single-track segments representing spacing, in miles, between adjacent sidings (Figure 5.3).
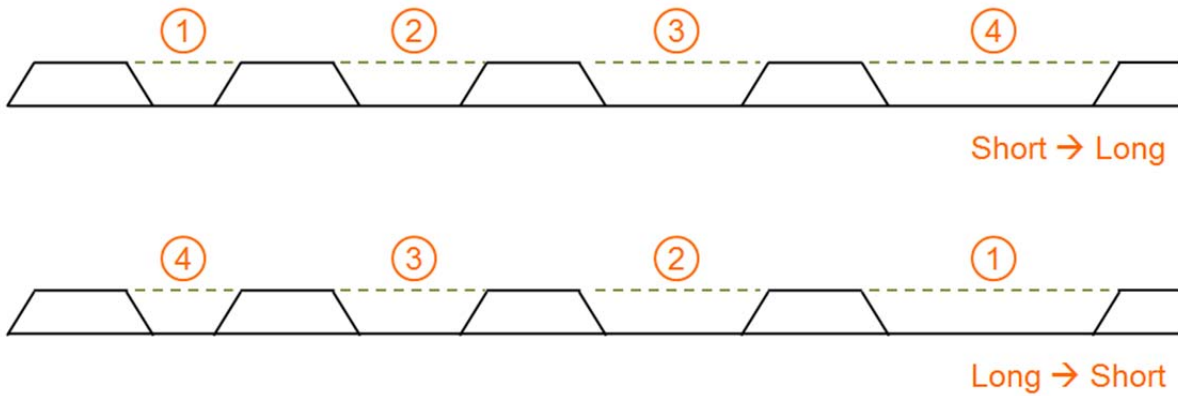
Figure 5.2: Generalized route with short-to-long and long-to-short build-out strategies; Circled numbers represent the order in which a siding connection is made

Figure 5.3: Initial 240-mile route layout for variable siding spacing and connection strategy experiment; Numbers represent the spacing (in miles) between adjacent sidings, which are represented by squares

The distances between sidings are arranged such that connections can be made in a balanced manner for both strategies under consideration. For example, consider the case of the *short-to-long* connection strategy. Initially, the sidings spaced at 8 miles in Section 1 and Section 3 are connected simultaneously, followed by the sidings spaced at 8 miles in Sections 2 and 4. This eliminates all of the bottlenecks between sidings spaced at 8 miles, leaving the shortest single-track sections as those between sidings spaced at 10 miles. The bottlenecks between

46

sidings spaced at 10 miles in Section 1 and Section 3 are then connected simultaneously, followed by the single-track segments between sidings spaced at 10 miles in Sections 2 and 4. This pattern will repeat itself incrementally until the longest single-track segments between sidings spaced at 16 miles are connected, and the entire route is composed of two-mainline track. The same procedure is followed for the *long-to-short* strategy, only the sequence is reversed so that sidings spaced at 16 miles are connected first, followed by 14, 12, etc.

The pattern described here is intended to experimentally isolate the effects of each build-out strategy. If the alternating pattern of building in Sections 1 and 3, and then 2 and 4 is not followed and a more random approach is taken, the route may end up unbalanced; one side of the route might be disproportionately double-tracked, while other segments remain sparsely connected. This could potentially confound the results, distracting from the goal of this study.

5.4 Results

After running simulations for the two experiments described in the previous section, train delay data were exported from RTC and used to define the incremental improvements in line capacity resulting from double-track installation. The results for each experiment are described in the following sections.

5.4.1 Impact of Initial Siding Spacing

Much like the Model 1 results obtained by Sogin et al. (2013a), Model 2 demonstrates a linear relationship with negative slope between percentage double track and delay (Figure 5.4). However, the delays are larger than those obtained for Model 1 with sidings spaced closer together (10 miles on-center as opposed to 16). This is not surprising, since it is expected that train delay in Model 2 will be greater than in Model 1 simply because there are longer bottleneck

47
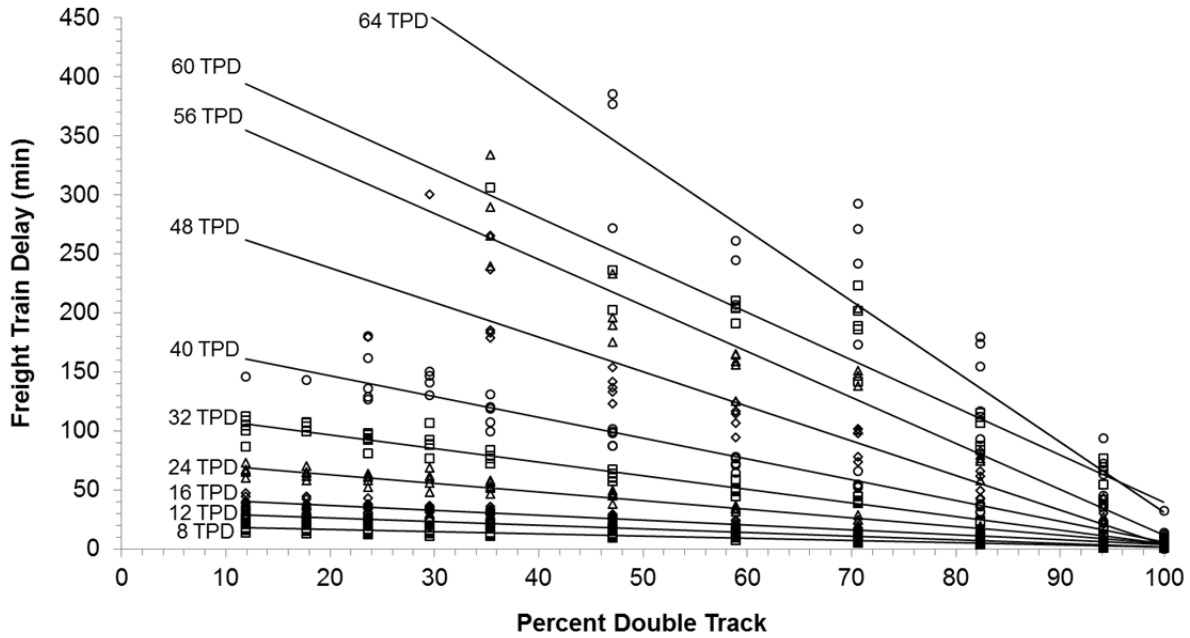
Figure 5.4: Train delay as a function of percent double track for a route with an initially even 16-mile siding spacing, with differing freight traffic volumes

sections throughout the route. The longer length of these single-track sections increases running time through the bottleneck, thereby reducing capacity and increasing delay.

The slopes of the regressed lines in Figure 5.4 provide additional information; there is a greater reduction in train delay (i.e. a steeper negative slope) for routes with higher traffic volumes. This, again, is expected, because routes with a higher density of train traffic also have more train meets so they experience greater congestion relief from additional second-mainline track. Both of these results increase confidence in the validity of the simulations.

Although linear delay reduction patterns are evident in Figure 5.4, there is more variability in delay for routes with higher train volumes and/or lower double-track percentages. This is due in part to a limitation of the software under congested operating conditions. If train delays are sufficiently high, RTC ends the simulation process for those scenarios since the train conflicts cannot be reasonably resolved. This is why there are no data points in the upper left of

Figure 5.4. If data could be collected at these higher volumes/lower double-track percentages, it might reveal non-linearity in this region.

A direct comparison of the results of Model 1 and Model 2 illustrates the influence of siding spacing on the delay response of double-track installation under equivalent traffic volumes (Figure 5.5). The lower line for each volume-pair represent the results from Model 1 (10-mile siding spacing), while the upper lines represent the results from Model 2 (16-mile siding spacing) shown in Figure 5.4. At 24 trains per day (TPD), the two lines are similar, indicating a roughly equivalent benefit from double-tracking, irrespective of initial siding spacing. However, in the case of 48 TPD, the two lines are much farther apart and the gap between them is disproportionately large compared to 24 TPD. For example, at 50 percent double track, the gap between the two 48 TPD lines is more than double the gap for the 24 TPD lines, even though the traffic volume is only twice as large. This indicates that siding spacing has a disproportionately
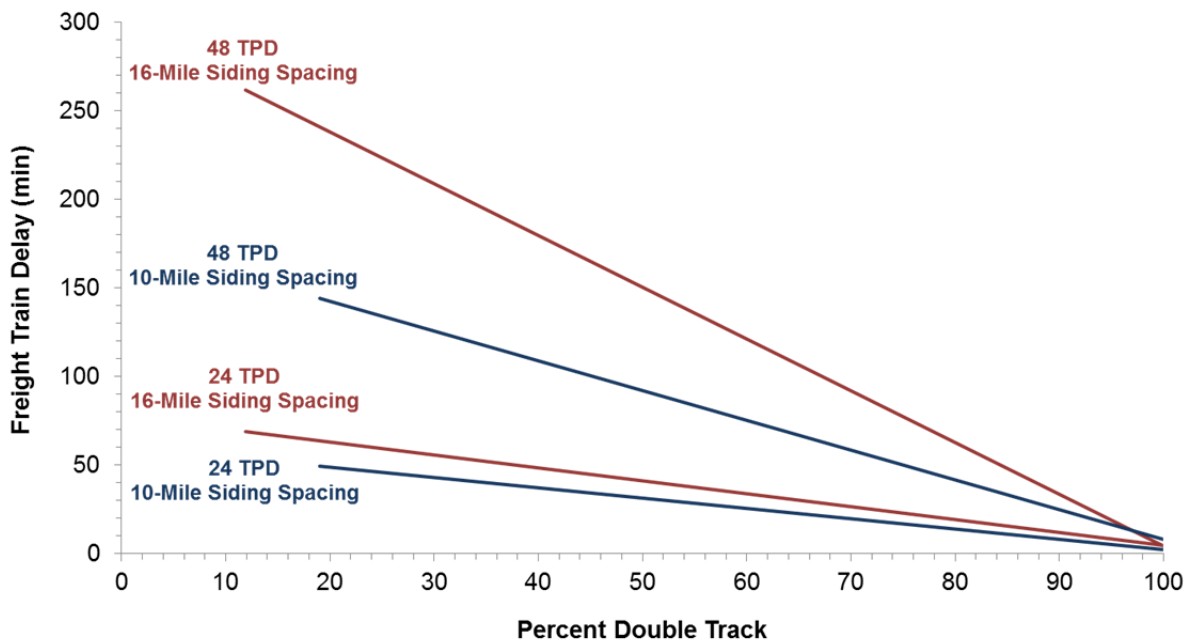
Figure 5.5: Train delay as a function of double-track percentage for two freight traffic volumes and two different initial siding arrangements

49

larger impact on delay for lines with higher traffic volume than for those with lower traffic volume. The relative slopes of the lines at 48 TPD also indicate that for the same high traffic volume, double track has a disproportionately greater benefit on lines with larger initial siding spacing. The difference in delay-response depending on the siding spacing distance provides additional motivation for the investigation of variable siding spacing conducted as the second part of this research. It also suggests that a *long-to-short* strategy might yield the best delay reduction response.

5.4.2 Variable Siding Spacing and Connection Strategy

As discussed above, one might expect that the *long-to-short* build-out strategy would provide higher initial incremental delay-reduction benefits since this strategy eliminates the longest bottleneck sections first. However, comparison of delay as a function of double-track percentage for the *short-to-long* and *long-to-short* build-out strategies (Figure 5.6) indicates this is not the case; the lines for each build-out strategy at equal train volumes almost entirely overlap one another. This result indicates that it does not matter whether longer-spaced sidings are connected first, or the opposite approach is used.

A more detailed quantitative look at the incremental double-tracking benefits of the *short-to-long* and *long-to-short* build-out strategies reveals subtler trends in the data (Table 5.2a and 5.2b). The incremental benefit of each step in the double-track construction process is calculated by taking the corresponding reduction in minutes of delay (per 100 train-miles) and dividing by the length of double track installed in that segment (expressed as a percent). The result is a measure of the rate of return on investment expressed in the units of minutes of delay-reduction per percent of double track installed (or minutes per %DT). By assigning specific
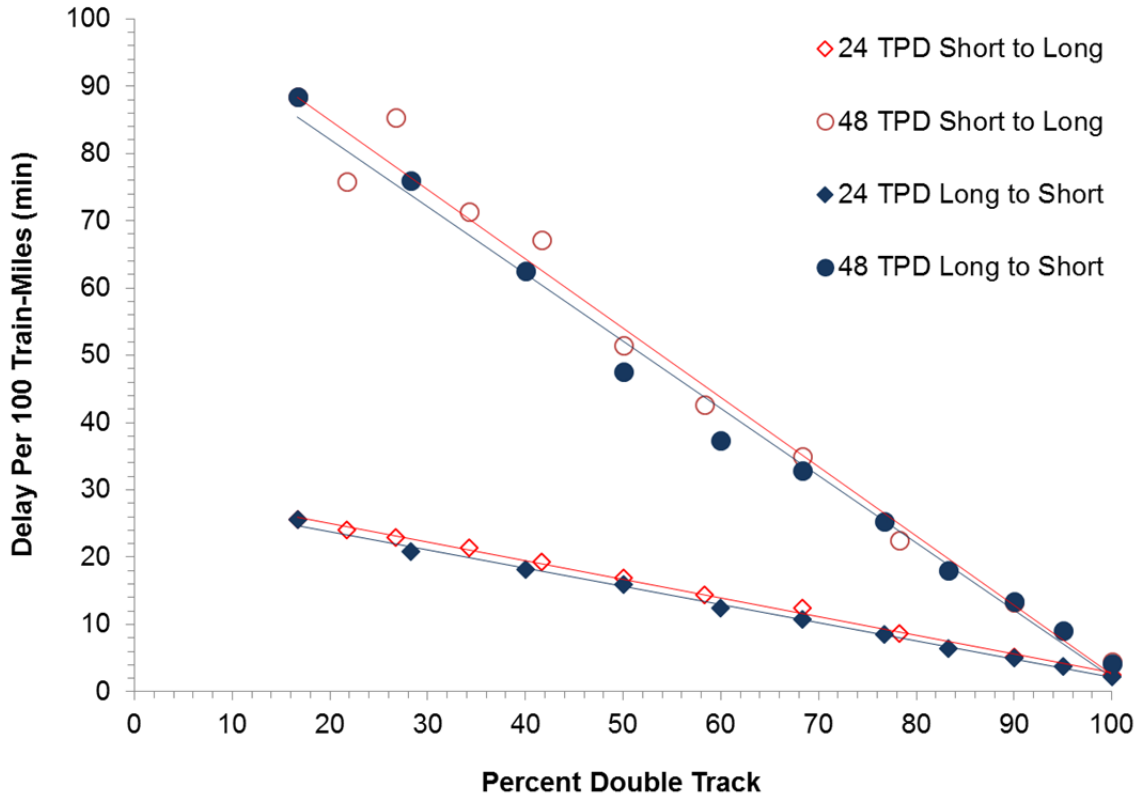
Figure 5.6: Delay per 100 train-miles as a function of percent double track for two freight volumes (24 TPD and 48 TPD) under two different build-out strategies

Table 5.2: Incremental delay benefits for the (a) *short-to-long* **and (b)** *long-to-short* **siding connection strategies**

Connection	Delay Reduction (minutes per % DT)	
	24 TPD	**48 TPD**
8-Mile Siding Spacing	0.26	0.31
10-Mile Siding Spacing	0.24	1.87
12-Mile Siding Spacing	0.29	1.88
14-Mile Siding Spacing	0.29	0.77
16-Mile Siding Spacing	0.28	0.83

(a)

Connection	Delay Reduction (minutes per % DT)	
	24 TPD	**48 TPD**
16-Mile Siding Spacing	0.32	1.11
14-Mile Siding Spacing	0.29	1.26
12-Mile Siding Spacing	0.24	0.72
10-Mile Siding Spacing	0.26	0.89
8-Mile Siding Spacing	0.28	0.94

(b)

dollar amounts to the cost of delay and cost of double-track installation per mile, this rate of return could be transformed into a benefit-cost ratio. In Tables 5.2a and 5.2b, the delay reduction values are based on averages for each unique siding spacing.

While the two tables appear similar, there are differences. For example, at 48 TPD and an 8-mile spacing, only 0.31 minutes per %DT are saved when these particular sections are connected first (i.e. *short-to-long* strategy), while 0.94 minutes per %DT are saved when they are connected last (i.e. *long-to-short* strategy). A reverse relationship in delay reduction rates is apparent for the 16-mile siding spacing segments, reaffirming the notion that connecting the longest bottleneck sections first yields a greater return on investment than connecting them last. Thus, the higher cost of eliminating longer bottleneck segments may be more easily justified if they are constructed earlier in the transition to double track when their rates of return tend to be higher.

A railroad starting the process of installing double track on this corridor would be likely to simulate the first set of connection alternatives and then evaluate the benefit-cost ratio of the different alternatives to select the first project. In this case for a traffic volume of 48 TPD, the rate of return for the 16-mile connections is 3.5 times the rate of return for the 8-mile connections. Thus, it is likely the *long-to-short* strategy would be adopted and the 16-mile connections built first. However, if the entire set of connections is simulated incrementally, connecting the 8-mile segments first via the *short-to-long* strategy allows the 10 and 12-mile segments to have a greater return than when they are used to make later connections in the *long-to-short* strategy. Thus, the incremental benefit of a double-track connection between sidings as measured by delay reduction per unit of double track installed is not purely a function of the length of the bottleneck segment. The rate of return for any individual incremental connection is

52

influenced by the size and number of different bottlenecks that have previously been eliminated and remain to be connected. The consequences of this finding will be more thoroughly considered in Chapter 6.

An example of this result is the role of the 12-mile siding connection. This connection is more important in the scenario where the longest bottleneck sections have not yet been double-tracked (i.e. *short-to-long* strategy), as opposed to if they already have been (i.e. *long-to-short* strategy). This is further illustrated by comparing the rate of return for segments in the uniform siding spacing cases (Figure 5.4) to the rate of return for segments of the same siding spacing distance in the non-uniform siding spacing cases. At the higher traffic volumes of 48 TPD, the delay reduction for the even 16-mile siding spacing model (Model 2) presented in Figure 5.4 is roughly 1.22 minutes per %DT, compared to the 0.83 minutes per %DT and 1.11 minutes per %DT shown in the non-uniform spacing models in Tables 5.2a and 5.2b. This indicates that these long, 16-mile connections have a larger delay benefit when part of an even, idealized line of many widely-spaced sidings, than for routes with non-uniform siding spacing and some shorter bottleneck segments.

5.5 Conclusions and Future Work

Routes with sparse sidings experience larger reductions in train delay (i.e. more congestion relief) via double-track installation compared to routes with sidings spaced closer together. Further comparisons revealed that siding spacing has a disproportionately larger impact on delay for lines with higher traffic volumes than for those with relatively lower volumes. This indicates that double-track installation offers disproportionately greater benefits on busy lines with larger initial siding spacing.

In regards to non-uniform siding spacing and connection strategies, the results showed that when the entire progression from single to double track is considered, there appears to be no difference in double-tracking longer bottleneck sections before shorter ones. The implication for railway applications is that the lowest-cost option (likely to be the connection of shorter-spaced sidings) should be the preferred option regardless of track infrastructure locations. The results did suggest that connecting the longest bottleneck sections first leads to the greatest initial return on investment in terms of reduction in train delay per unit of double track installed. However, these results were inconclusive, and suggest that more research is warranted regarding double-track installation strategies.

Based on the limited data obtained via RTC for scenarios where relatively high freight traffic volumes were combined with low double-track percentages, the relationship between percent double track and train delay appeared to be linear. However, due to constraints in the simulation software, the most extreme conditions could not be simulated.

The results for non-uniform siding spacing connection strategies that contrast with expectations (i.e. connecting longer bottlenecks first provides the greatest return on investment) could be clarified via experimentation where the number of siding lengths considered on the route in Figure 5.3 is reduced from five to two. More specifically, if only 8- and 16-mile siding spacings were considered and the intermediate spacings eliminated to focus solely on the two extremes, the results might provide a sharper contrast between the *short-to-long* and *long-to-short* siding connection strategies.

A zonal demand model could also be used instead of the two strategies presented in this study. This model would not follow a predetermined order of connection projects; rather, it would incorporate a check of cumulative delay at each point in the route for each simulation.

This delay would then be used to determine where along the route train delays are concentrated. Installation of double track in those sections would become the next incremental expansion projects selected for implementation. This process would be iterated after every route simulation in RTC, and would therefore represent an evolving, real-time decision strategy for double-track installation, as opposed to the two predetermined strategies used in this study. This strategy could then be compared against the others in order to determine an optimal, streamlined process for identifying the projects with the most potential for further engineering evaluation.

Finally, the results in this chapter were obtained for the case of homogenous freight traffic. A mixed-use corridor with freight and passenger trains operating at different speeds, with the consequent need for train passes as well as meets, would likely lead to different relationships between double-track installation and train delay. It is possible that in this situation the results of *short-to*-long and *long-to*-short siding connection strategies may differ.

CHAPTER 6
SINGLE TO DOUBLE TRACK: SIDING CONNECTION
LENGTH, POSITION, AND ORDER

An earlier version of this research appears in:

Atanassov, I. & C.T. Dick. 2015c. Influence of siding connection length, position, and order on the incremental capacity of transitioning from single to double track. In: Proceedings of the 94th Annual Meeting of the Transportation Research Board of the National Academies, Washington, D.C.

6.1 Introduction

The railway infrastructure in the United States is primarily composed of single-track mainlines with limited capacity to maintain required levels of service as traffic volumes grow and operational complexity increases (Cambridge Systematics 2007). Where freight and passenger services share infrastructure, increasing demand for both types of rail transportation can have a compounding effect (Dingler et al. 2012, Dingler et al. 2013). On these shared rail corridors, the need to expand infrastructure to avoid congestion and mitigate delay to freight and passenger trains will happen at lower traffic levels than when traffic is more homogeneous. While there are several approaches to increasing rail line capacity, the primary infrastructure expansion strategies involve extension of existing passing sidings to accommodate meets between three trains, or construction of additional passing sidings (Shih et al. 2014). While these steps may provide initial solutions to the problem, it will eventually become necessary to consider installation of double-track segments to ensure capacity for future rail traffic volumes (Sogin et al. 2013b).

Network-level models can help railroad practitioners identify the routes where double-track construction will most effectively increase capacity (Lai & Barkan 2011). However, capital program planners still face the complex task of selecting among numerous candidate segments within each critical route. While engineering obstacles such as tunnels and large

56

bridges will eliminate some locations, and local operating needs such as switching work, yards, and grades may make double track more attractive on certain segments, planners are still faced with the daunting task of selecting between a large number of project alternatives. Detailed simulation and engineering investigation to establish the cost and benefit of all options is impractical due to time and resource constraints. Thus, railroad planners often use simple heuristics to screen the alternatives and select a smaller subset of double-track projects for detailed evaluation. In discussion with Class 1 railroad planners, examples of double-track heuristics developed through experience include:

- Make longer double-track connections first, followed by shorter connections

- Double track offers the greatest return on segments approaching terminals

- Locations corresponding to frequent train meets are ideal candidates for double track

- Initial double-track segments offer little benefit until they are connected by additional segments and the benefits compound; thus it is better to continue adding double-track segments along a route, as opposed to installing the first segment on a different route

Overall, these heuristics suggest that connection length, route position, and order can serve as quick indicators of the potential incremental capacity offered by installation of double track between a pair of existing passing sidings.

The analyses that follow aim to determine if the incremental delay-benefits of installing segments of double track on single-track corridors exhibit trends in connection length, route position, and order that correspond to the above heuristics. If distinct trends are discovered that support the above heuristics, or alternatively suggest different rules, the results can serve as a

guideline for a more streamlined decision-making process by helping to quickly identify the types of projects with the most to gain from double track. With a smaller number of prioritized project alternatives, railroads can better utilize their modeling, planning, and engineering resources in conducting a more detailed analysis to make final project selections.

To complement previous research results pertaining to lines with variable siding spacing and homogeneous traffic (Chapter 5), this study introduces modified spacings and heterogeneous traffic to further investigate the connection-length heuristic.

6.2 Background

Measurement of rail capacity and delay characteristics for single-track mainlines has been well-covered in existing literature, and such research has been extended into studies on the delay benefits of double-track installation. Mitra et al. (2010) introduced parametric methods for the estimation of single-track railway capacity. Lindfeldt (2010) broadened the physical scope of single track analysis to consider and analyze the operational dynamics inherent to a double-track rail corridor configuration. Gussow and Welch (1986) analyzed the capacity of partial double-track lines, and the effect of track-infrastructure distribution on system performance. The subsequent analyses presented in this chapter, however, are rooted in results obtained in Chapter 5 as well as the work done by Sogin et al. (2013a). Results showed that for idealized corridors with even 10- and 16-mile siding spacing, double-track installation provided a linear reduction in train delay for differing levels of freight traffic (Figure 5.5). The reduction in delay resulting from double-track installation is consistent with previous findings that identified train meets as primary causes of delay, with double track allowing for a larger percentage of trains to avoid meets altogether (Dingler et al. 2010).

Recalling the characteristics of the lines in Figure 5.5, the slopes for the 16-mile study (representing minutes of delay reduction per percent double track installed) are steeper than their corresponding slope in the 10-mile study. On a macro scale, this phenomenon suggests that, given a particular traffic volume on a line with non-uniform siding spacing, the connection of longer-spaced sidings (or elimination of the longest single-track bottlenecks) should reduce delay by an amount greater than is achievable via connection of shorter-spaced sidings (or elimination of the shortest single-track bottlenecks). This hypothetical effect is visualized by two contrasting delay responses for an arbitrary route consisting of non-uniform siding spacing (Figure 6.1). The lower trajectory depicts a scenario where longer-spaced sidings are connected first. Train delay is hypothesized to exhibit an initial sharp decline followed by a reduction in incremental benefit resulting from the connection of remaining shorter-spaced sidings. The upper curve illustrates the opposite response where shorter siding connections are given initial priority and increasing returns are observed as the final long connections are made.
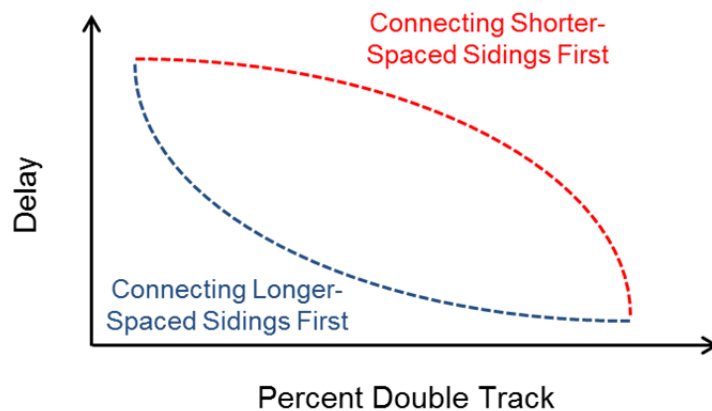
Figure 6.1: Hypothetical delay response curves for two different siding connection strategies

These hypothetical response curves represent conventional industry heuristics. The head of service design at one Class I railroad in the United States favors an incremental approach of gradually adding double track between sidings, connecting the longest intervening sections first, then proceeding to connect the shorter ones (Martland 2008). The research presented here investigates the sensitivity of incremental line capacity to different double-tracking strategies in order to distinguish between the two hypotheses. In so doing it should help railroads understand the validity of the standard "long-to-short" heuristic.

6.3 Methodology

In practice, it is the running time between sidings and not the siding spacing distance that controls the capacity of a single-track line. However, for this study, the maximum track speed on all sections of the hypothetical line is equal and the grade is also uniform, resulting in uniform operating speeds along the route (50 mph freight trains, and 110 mph passenger trains). Thus, the distance between passing siding centers can be used as a direct proxy for the running time between sidings. Part of the following methodology is carried over from the previous chapter, and is modified and extended through additional steps.

There are numerous strategies that can be employed when selecting the order of existing sidings to connect into double-track sections on a route with non-uniform siding spacing. Again, heuristics imply that sidings spaced farthest apart should be connected first. Such a strategy ensures that the longest bottleneck sections are removed from the route, presumably leading to the highest potential reduction in train delay. To test this heuristic and to provide the greatest potential contrast in delay response, the same two build-out strategies from Chapter 5 (i.e. *short-to-long* and *long-to-short*) are used but on a different initial 240-mile route (Figure 6.2).
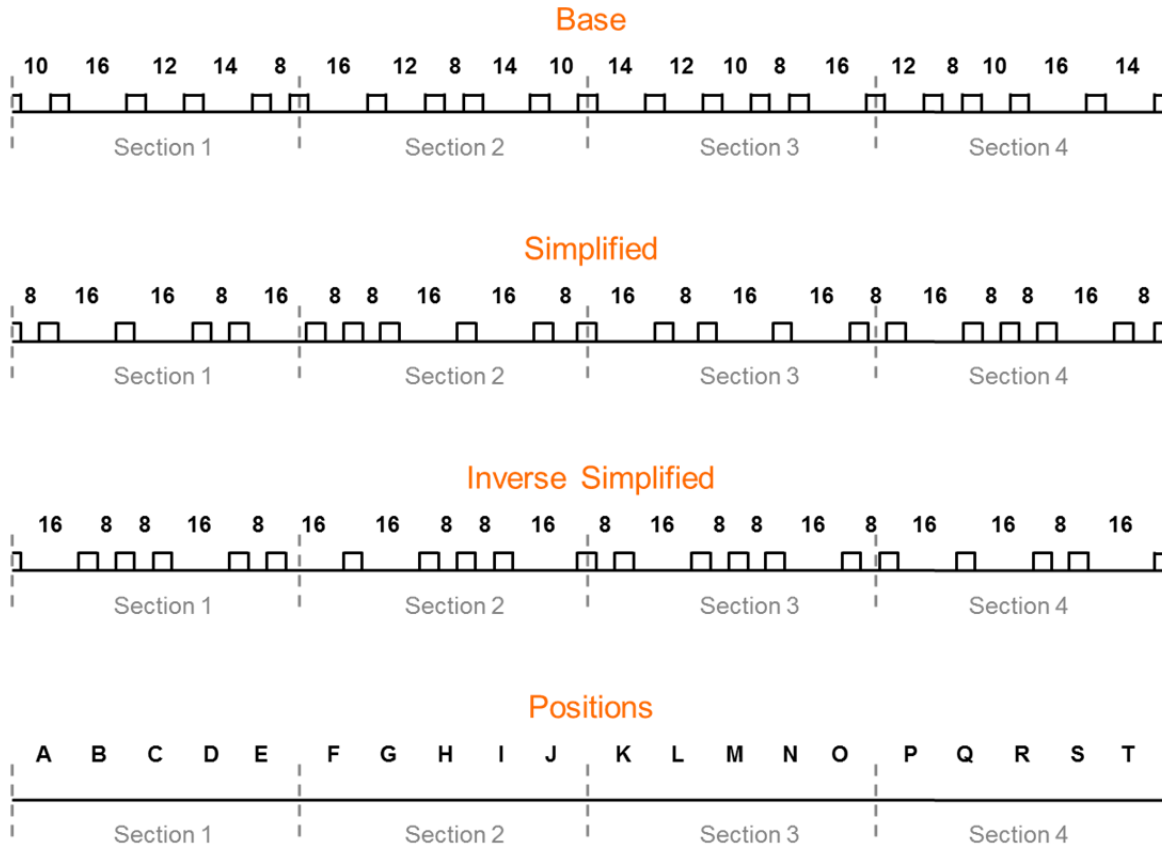
Figure 6.2: Initial siding arrangements for the 240-mile route; Squares indicate sidings and numbers indicate the distance (in miles) between them; Bottom graphic depicts the naming convention for relative siding locations

The numbers in Figure 6.2 represent the spacing, in miles, between adjacent sidings and lead to particular connection patterns on the *base*, *simplified* and *inverse simplified* layouts. An example of the connection pattern process is provided in the Methodology section of Chapter 5 (Figure 5.2).

The *base* layout is the same layout that was analyzed in Chapter 5. The *simplified* layout shown here, however, is a modification of the *base* arrangement that only uses two siding spacings (8 and 16 miles), as opposed to the original five. The purpose of focusing on these two extreme siding lengths is to potentially show a sharper contrast between the delay response of the two connection-order strategies (*long-to-short* and *short-to-long*). In the *simplified* layout,

however, siding connections are no longer made in pairs, as in Chapter 5, but rather one-by-one. Even though connections are no longer paired, successive connections still follow the general pattern of connecting in Sections 1 and 3, followed by Sections 2 and 4 – the reasons for which are described in Chapter 5.

The *inverse simplified* layout was created to isolate the influence of route position, as opposed to connection length, of the new double-track segment. More specifically, where 8 miles exist between sidings in the *simplified* model, 16 miles exist in the *inverse simplified* scenario, and vice versa. Simulation and observation of this *inverse simplified* scenario in comparison to the *simplified* layout can determine if the delay-reduction observed for a particular connection is a function of the siding connection length, the position of the siding connection along the route (e.g. near the middle of the route, close to terminals, etc.), or some combination of both of these factors.

6.4 Results

After running simulations on the three siding connection arrangements described in the previous section, delay data were exported from RTC and used to characterize the relationship between train delay and double-tracking strategy. The results for each experiment are detailed in the following sections.

6.4.1 Base Scenario with Range of Connection Lengths

Simulation of the *base* scenario route with its range of siding spacing distances is carried out for both homogeneous and heterogeneous traffic mixtures. The homogeneous-freight-traffic scenario (Figure 6.3a) was developed in Chapter 5, and is presented again to compare to the new heterogeneous case (Figure 6.3b). 'TPD' is an abbreviation of Trains Per Day.
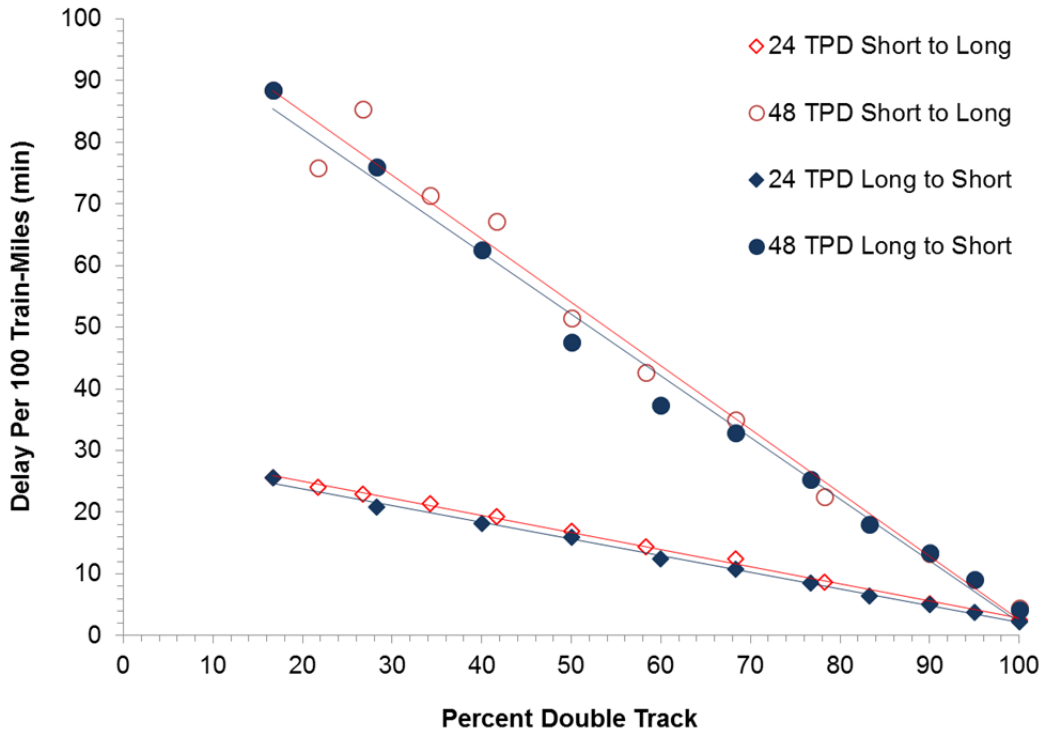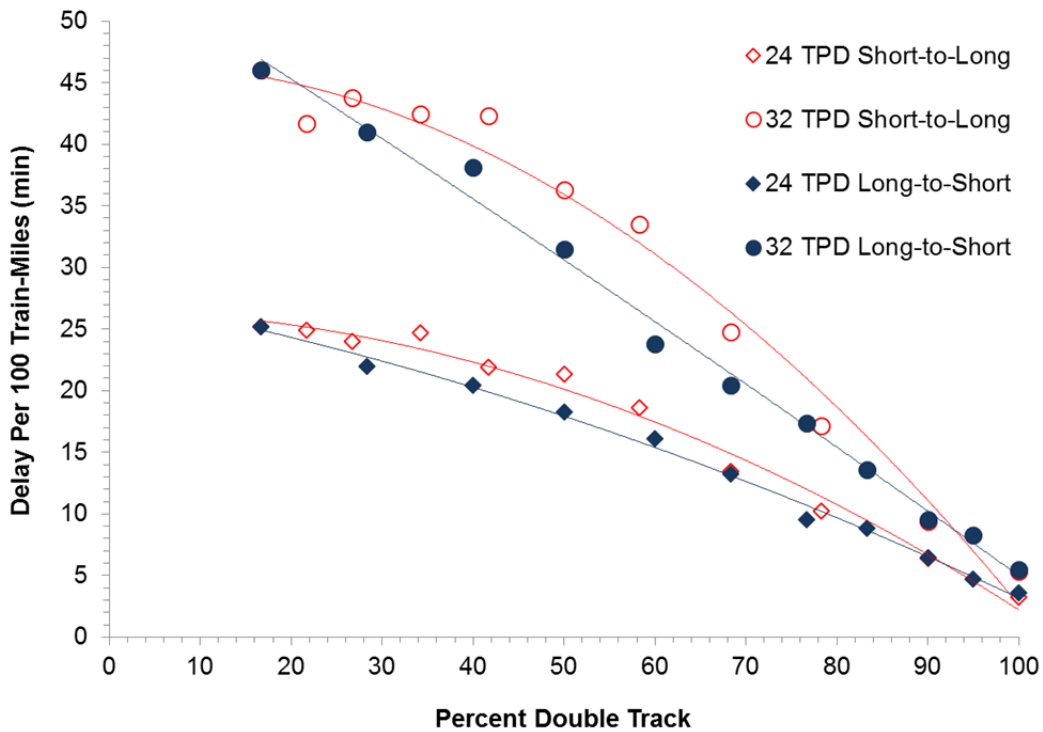
Figure 6.3: Delay as a function of percent double track for the (a) homogeneous and (b) heterogeneous case of the *base* scenario; Heterogeneous case includes 75% freight and 25% passenger

Compared to the hypothetical delay response (Figure 6.1) characterized by inversely curved delay-trajectories dependent on the type of connection strategy being employed, the simulated results show little difference between the *short-to-long* and *long-to-short* connection strategies (Figure 6.3a and 6.3b). There is almost no difference in the linear trends of each connection strategy for the case of homogeneous freight traffic. In the heterogeneous case, the introduction of priority passenger trains causes some separation between the delay-response of the two strategies. While delay reductions for the *long-to-short* connection patterns at both traffic volumes in the heterogeneous case show a fairly linear trend, the *short-to-long* connections are somewhat more curvilinear, reminiscent of the curves in Figure 6.1. More specifically, delay values remain relatively static for lower percentages of double track (i.e. when shorter double-track sections are being added), and only begin to drop off significantly near the 50 percent double track mark. This point corresponds to when longer double-track connections (12-or-more miles) are starting to be made. This pattern of delay reduction is comparable to that of the upper hypothetical curve in Figure 6.1.

In the case of homogeneous freight traffic at 24 and 48 TPD, there is no indication of any substantial benefit resulting from connecting longer bottleneck sections first (Figure 6.3a). This suggests that the lowest-cost option (likely to be the connection of shorter-spaced sidings) should be preferred regardless of infrastructure location (Shih et al. 2014). Inspection of the trends in the heterogeneous case, however, suggest there is an increased delay-benefit to connecting longer bottlenecks first. In particular, connecting shorter bottlenecks first does little to reduce delay until sizable amounts of double track have already been installed and longer connections have been made. These trends parallel some of the simple heuristics described earlier in this chapter.

6.4.2 Simplified Scenarios with 8- and 16-Mile Connections

A potential limitation of the *base* scenario is that it involves a broad range of siding spacing distances. Having multiple siding spacing distances (and not a lot of difference between them) may actually hinder the investigation into the relative effects of the *long-to-short* and *short-to-long* connection strategies. For example, there may not be much difference between making 12- and 14-mile connections. This could be an explanation for the lack of separation in the delay response observed in Figure 6.3a and 6.3b. Therefore, in order to provide greater contrast in the lengths of connections being made by the *long-to-short* and *short-to-long* strategies, the experiment was repeated for heterogeneous traffic on the *simplified* and *inverse simplified* layouts. As mentioned previously, the *simplified* and *inverse simplified* scenarios drop three of the intermediate siding connection lengths (10, 12, and 14 miles), leaving only the two "extremes" of 8 and 16 miles.

Plotting the results from simulations, it is apparent that the actual delay-response recorded for both route arrangements (*simplified* and *inverse simplified*) does not resemble the hypothetical response predicted earlier (Figure 6.4a and 6.4b). For both route arrangements, the *long-to-short* and *short-to-long* curves overlap, and even intertwine, as the transition to full two-mainline track progresses. The *inverse simplified* scenario does exhibit a slight separation between the *long-to-short* and *short-to-long* curves but the effect is small. These results, at least when presented in graphical form, do not support the simple heuristic that connecting longer sections first will result in larger delay reductions than when prioritizing shorter connections.
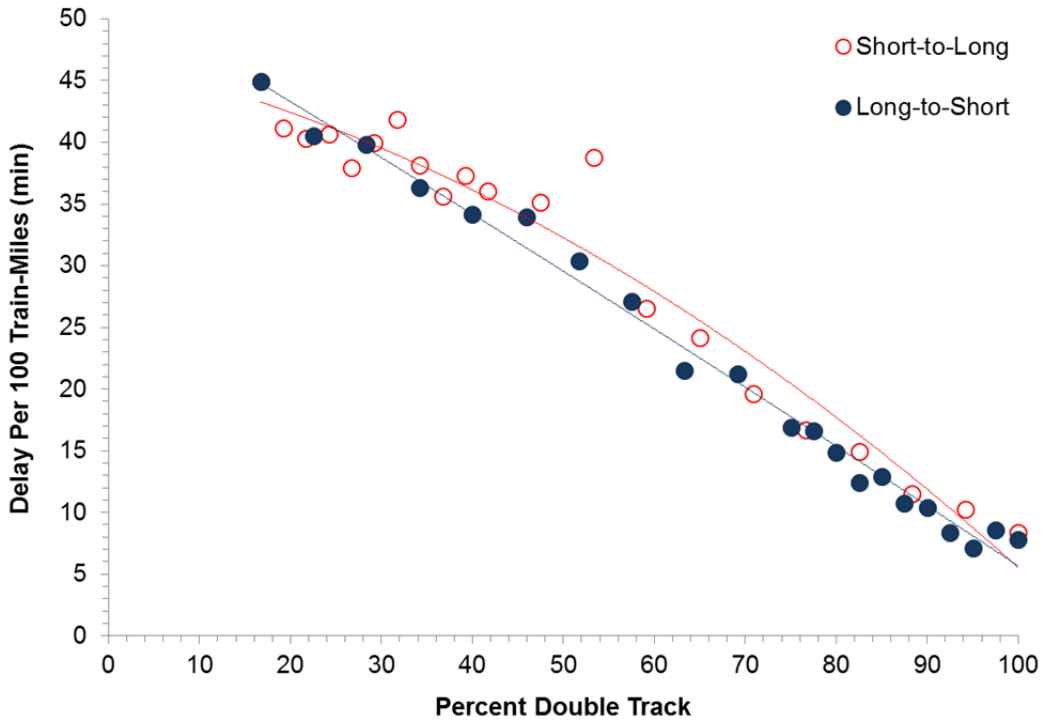
(a)
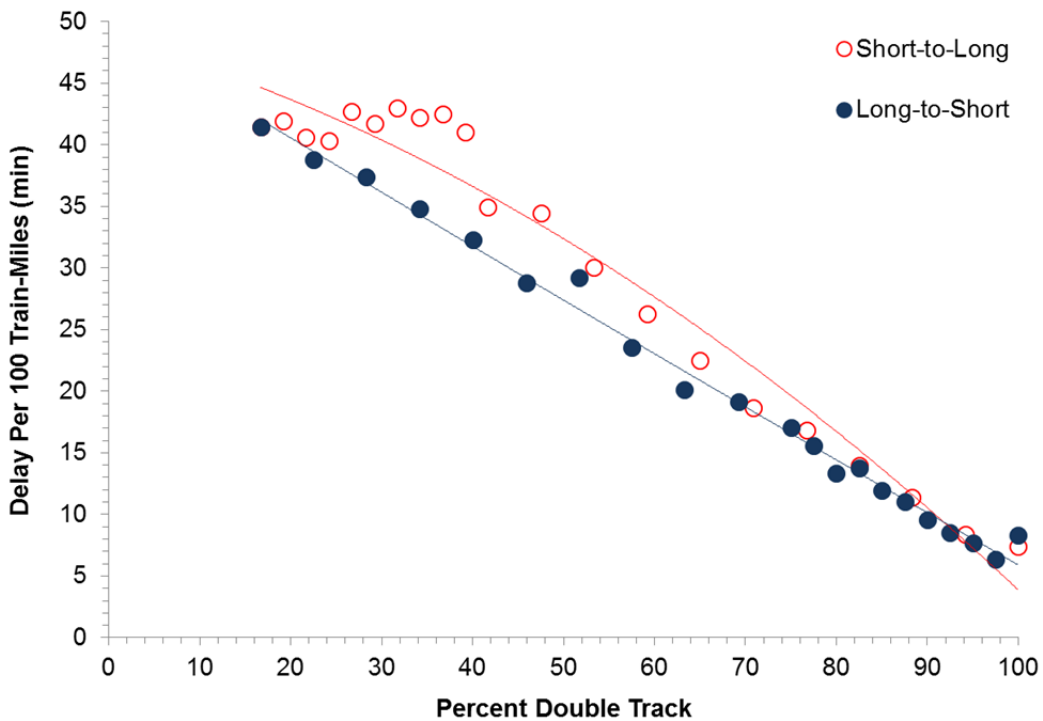
(b)

Figure 6.4: Delay as a function of percent double track for the (a) *Simplified* **and (b)** *Inverse Simplified* **scenarios; Results shown are for 32 TPD, 75% freight and 25% passenger**

6.4.3 Effect of Siding Connection Location

For a given route location, the four combinations of the *simplified* and *inverse simplified* routes and the *long-to-short* and *short-to-long* connection orders provide results for four distinct connection-project circumstances:

1. Segment is short (8 miles) and connected early in the progression to double track

2. Segment is short and connected late in the progression

3. Segment is long (16 miles) and connected early in the progression

4. Segment is long and connected late in the progression

These four different results are summarized graphically for each route position (Figure 6.5). For the sake of comparing across different connection lengths, the delay values are normalized by the length of double track installed to make each siding connection. This process takes the corresponding reduction in minutes of delay per 100 train-miles for each new double-track segment, and divides by the length of the double track installed for that connection (expressed as a percent). The result is a measure of the rate of return on investment expressed in the units of minutes of delay reduction per percent of double track installed, or minutes per percent double track. If the heuristic of making long connections first is to hold true, the third bar at each position, corresponding to the condition of "long connected early", should show the largest delay reduction. It is apparent, however, that this is only the case for a small number of route positions and is not a general trend.

Figure 6.5 also illustrates the influence of double-track positioning on line capacity. If the route-position of a double-track connection has no influence on delay, then all the bars for a given length/order should be at or around the same height. The magnitude of the bars, however, shows considerable variability. Certain route locations do provide a consistently larger delay
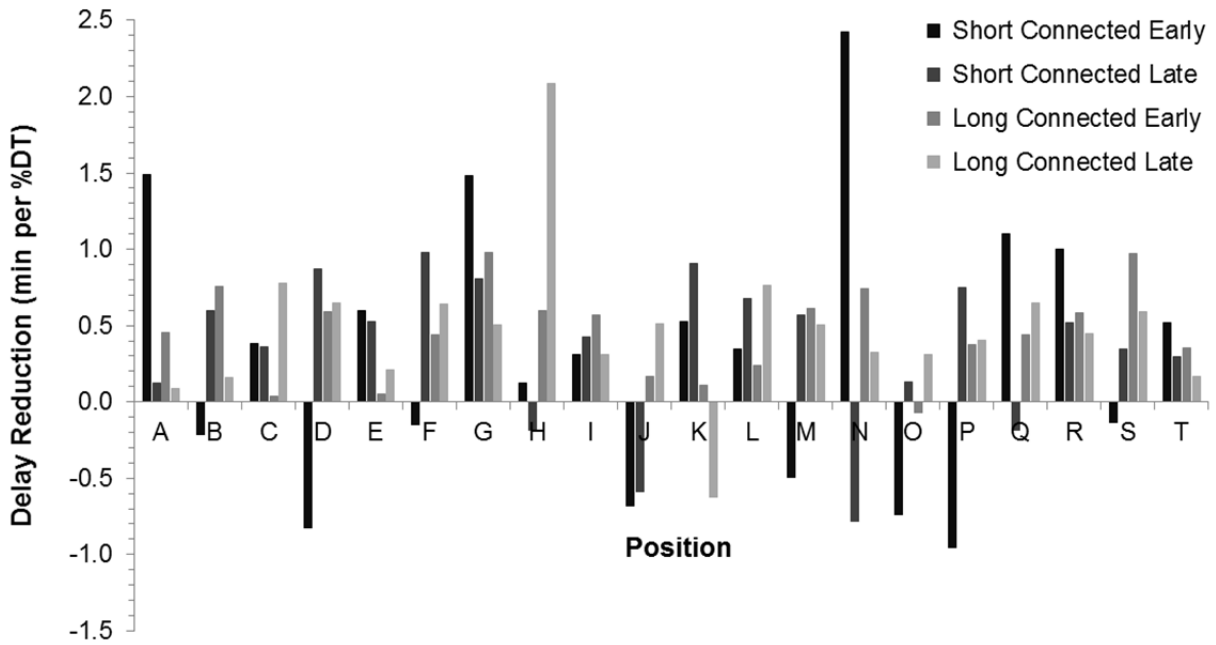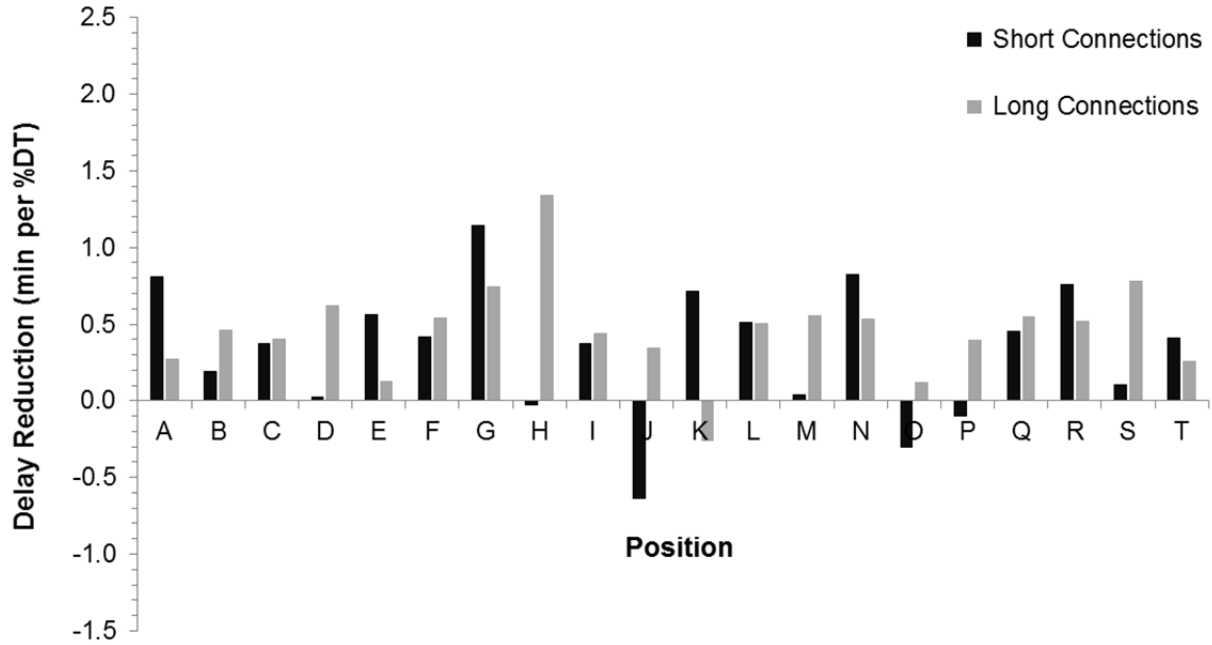
Figure 6.5: Average delay reduction for each combination of siding connection position, arrangement (i.e. *Simplified* and *Inverse Simplified*), and connection type; Results are for 32 TPD, 75% freight and 25% passenger

reduction than others; however, there is no obvious structure to this response linking these segments to specific route features (e.g. middle of the route, near terminals, etc.). Thus, position may not be the most useful heuristic on its own. Rather, there is an interaction between double-track position, connection, and order.

This finding is further supported by examining the relative delay-reduction of the four different circumstances at each particular position. Although some positions show relatively consistent delay response for each of the four circumstances, most show wide variation for different combinations of connection length and order. Overall, this comparison suggests that the length of the single-track bottleneck segment should not be the sole consideration in prioritizing projects; certain route positions may offer a greater return on investment. This may also help explain why the results in Figures 6.3 and 6.4 do not reflect the hypothesized relationship established earlier.

68

The three variables (length, position, and order) in Figure 6.5 can be separated out to show simpler, two-variable interactions with connection length graphed as a function of position (Figure 6.6) and connection order graphed as a function of position (Figure 6.7). In Figure 6.6, delay values for "short connected early" and "short connected late" are averaged for each position, and compared to their long-connection counterparts, averaged in a similar manner. This isolates the interaction between connection length and position. In Figure 6.7, delay values for "short connected early" and "long connected early" are averaged for each position, and compared to their late-connection equivalents. This isolates the interaction between connection order and position.

If position has no significance on delay reduction, then all the bars within a given series in either Figure 6.6 or 6.7 should be uniform. From the figures it is evident that this is not the case, suggesting that certain positions may provide larger delay reductions than others. In Figure 6.6, if connection length is not important, then the two bars at each position would be uniform. The bars in the figure show that this is not the case; pairs of bars at each position often have very different values. Similarly, if connection order is not important, the two bars at each position in Figure 6.7 would be uniform. Again, pairs of bars in the figure are non-uniform. These results support the notion that connection length, position, and order must all be considered together when prioritizing projects.

Figure 6.6: Average delay reduction for short and long connections made at each position; Results are for 32 TPD, 75% freight and 25% passenger
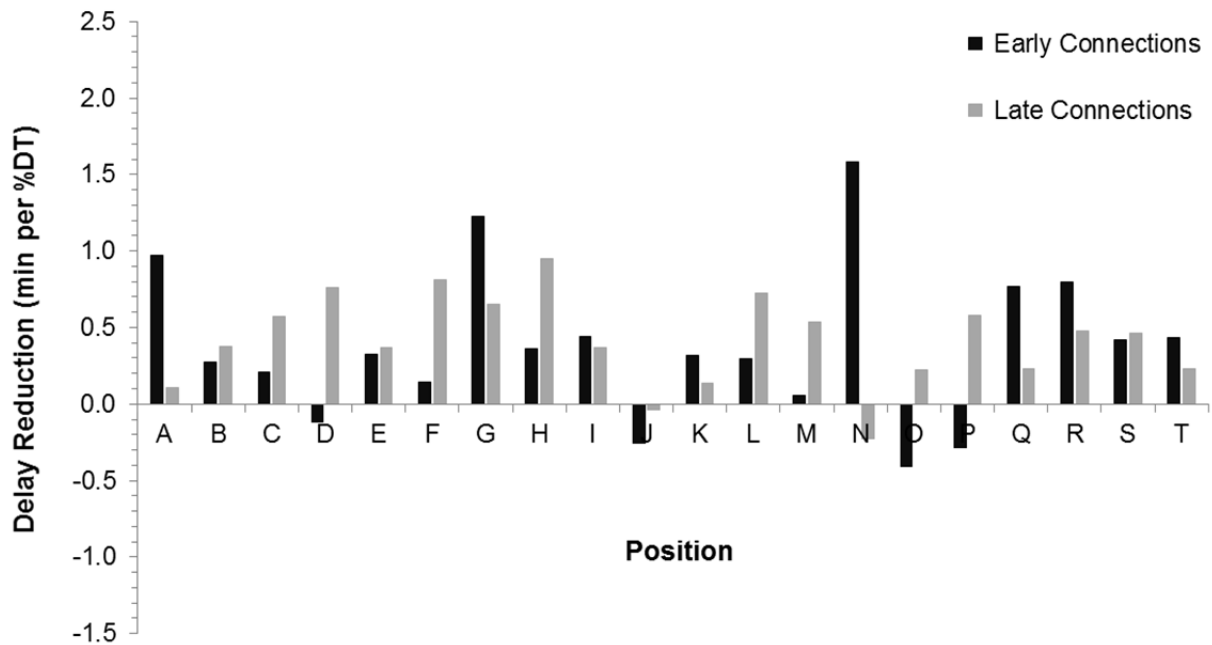
Figure 6.7: Average delay reduction for early and late connections made at each position; Results are for 32 TPD, 75% freight and 25% passenger

6.4.4 Effect of Siding Connection Order

A further reorganization of the simulation data was used to investigate the role of siding connection order on delay reduction (Figure 6.8). Each data point is sequentially associated with the order a connection project was completed over the complete progression from single to double track, regardless of connection length or position along the route. For example, the delay of all projects completed as the fifth step in the progression are averaged together to create the data point for Step Five.
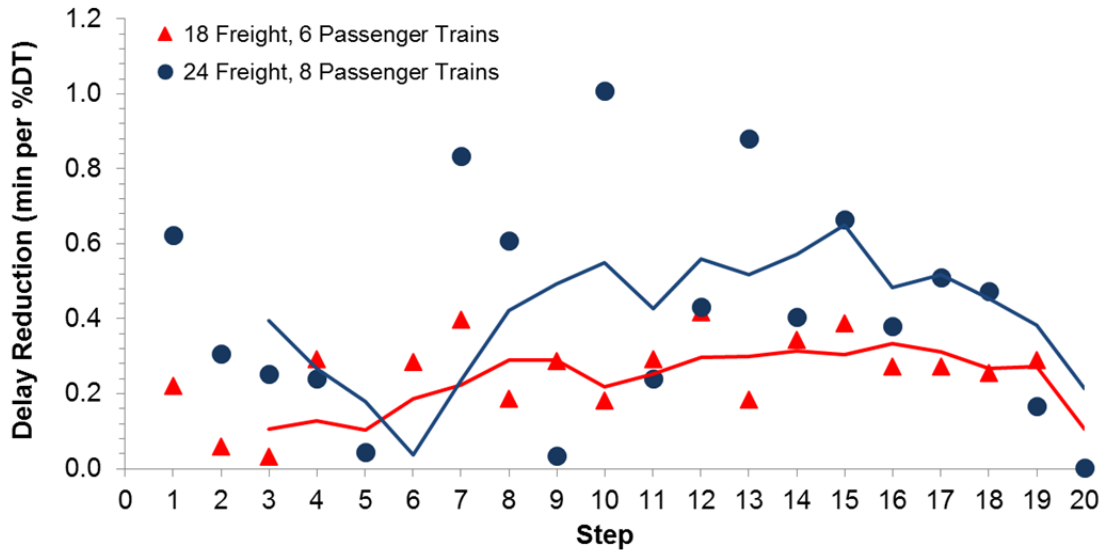
Figure 6.8: Delay reduction as a function of siding connection order (with 3-step moving average)

A 3-step moving average for each traffic volume is included to bring order to the highly variable distribution of average delay values. Note that projects ordered in the latter half of the double-tracking progression typically show higher delay-reduction values as compared to projects completed near the beginning or very end of the progression. This finding suggests that there are some economies of scale to adding double-track connections in that later connections compound the benefits of previous connections. While this supports the initial order-heuristic

71

described in Section 6.1, the weak trend suggests that connection order is not a dominant decision factor. Thus, order should be factored into track-expansion decision-making in conjunction with length and position.

6.4.5 Comprehensive Results

A primary objective of this study was to determine if there is significant delay-benefit in connecting longer-spaced sidings first, as opposed to shorter connections. Combining the simulated delay results across all three route layouts (*base*, *simplified*, and *inverse simplified*), the effects of siding connection length on line capacity can be quantified (Figure 6.9). Again, delay values are normalized by the length of double track installed, enabling comparisons between the two siding-connection lengths.

The larger normalized average delay-reduction values for 16-mile siding connections compared to 8-mile connections are evident, consistent with the benefit of prioritizing longer bottleneck sections for initial double-tracking. The values presented here suggest that longer connections are approximately 50 percent more effective at reducing delay as opposed to shorter
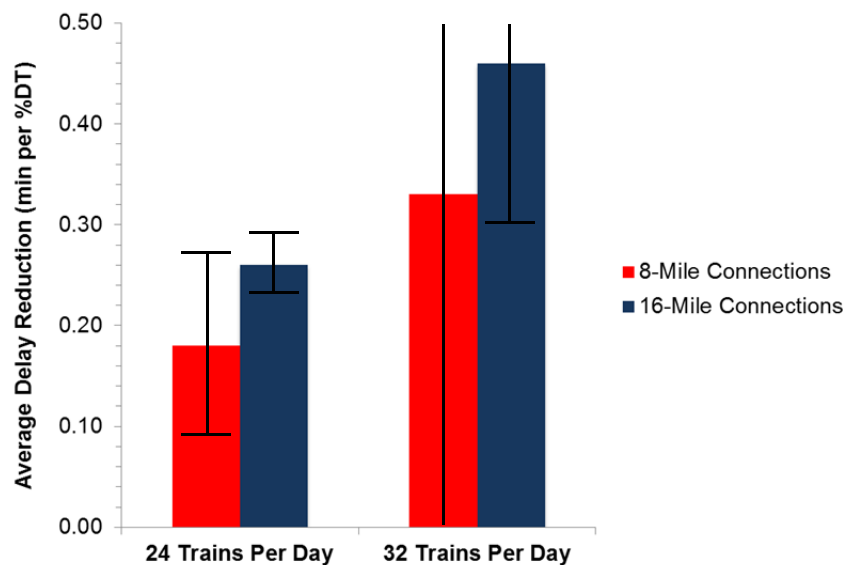
Figure 6.9: Summary of average delay reduction (with overlaid variance bars) for the connection of 8- and 16-mile siding spacings

72

connections. The delay variance values for 16-mile connection are substantially less than the 8-mile projects. The difference in variance indicates that the longer connections provide more consistent delay reduction, while shorter connections are more sensitive to the effects of route position and connection order.

6.5 Conclusions

Initial double-track project alternatives are often identified using simple heuristic rules regarding connection length, position, and order. Analysis of different siding connection strategies on a corridor with non-uniform siding spacing did not clearly support any of the heuristic approaches as the definitive rule for placement of new double track sections. The results demonstrate that the delay response of siding connection projects is influenced not only by the length of the connection being made, but by its position along the route, as well as the order that these connections are made within the full progression from single to double track. In particular, double-tracking projects completed in the latter half of the entire progression from single to double track appear to have a greater delay-based return on investment. While longer connections appear to provide more consistent delay reduction, shorter connections are more sensitive to the effects of route position and connection order, and can provide substantial delay reductions under the right conditions. These findings suggest a more holistic planning approach with more complex heuristics, requiring factor combinations of connection length, order, and position in order to properly support the initial screening of double-track project alternatives. When developed, a more comprehensive set of heuristics will lend themselves to practitioner applications in the form of a more efficient and effective decision-making process for capital expansion projects.

CHAPTER 7
INCREMENTAL CAPACITY IN TRANSITIONING FROM DOUBLE TO TRIPLE TRACK

An earlier version of this research appears in:

Atanassov, I. & C.T. Dick. 2015d. Incremental Capacity in transitioning from double to triple track on shared rail corridors. In: Proceedings of the International Association of Railway Operations Research (IAROR) 6th International Seminar on Railway Operations Modelling and Analysis, Tokyo, Japan.

7.1 Introduction

A substantial portion of the North American railroad network consists of single-track mainline with passing sidings or short segments of double track. These double-track segments often include routes where freight trains share trackage with passenger trains. Anticipated freight traffic growth and the expansion of commuter rail services on mainlines serving major urban areas suggests that further capacity upgrades will be needed. One option for achieving this is the addition of a third track on existing double-track corridors. Consideration of this investment is reinforced by the higher speed and frequency of passenger train operations on freight-dominated rail lines. Cambridge Systematics (2007) presented information regarding the capacity loss resulting from the operation of multiple train types (Table 7.1). Martland (2008) cites one North

Table 7.1: Practical capacity of rail lines with a differing number of tracks; CTC or TCS control (Cambridge Systematics 2007)

Number of Tracks	Trains Per Day	
	Practical Maximum with Single Train Type	Practical Maximum with Multiple Train Types
1	48	30
2	100	75
3	163	133

American Class I Railroad as considering that triple track is needed if traffic volume reaches 100 trains per day, and steep grades, slow speeds, or passenger operations are present.

Previous research on North American applications of partial triple-track installation has focused on qualitative discussion of the capacity benefits of this type of infrastructure expansion. More specifically, Tobias et al. (2010) investigated, via simulation models, the inability of double-track corridors to provide sufficient capacity to sustain the expected 20-year passenger and freight traffic growth along a particular shared-use rail corridor in the United States. Double track, while allowing simultaneous operation of freight and passenger trains, led to an unacceptable reduction in train speeds, an increase in delays, subpar on-time performance, and poor resiliency to recover from disruptions. Tobias forecasted the need for triple-track installation to deal with these operational maladies.

The research presented in this study seeks to characterize the relationship between incremental line-capacity and the phased transition from double to triple track. This characterization considers the overall length of triple track along a route and also how turnout arrangement at crossovers influences capacity.

Previous research conducted by Lindfeldt (2012) and Sogin et al. (2013a), and the research presented in Chapter 5, revealed that for idealized single-track corridors with uniform siding spacing, double-track installation provided a linear reduction in train delay across a wide range of freight traffic volumes. The study described in this chapter seeks to determine if the train-delay response associated with third-mainline construction follows this same linear trend in delay reduction.

Incremental construction of a third-mainline track is substantially more complex than its single-to-double-track counterpart due to the added consideration of crossover locations, turnout arrangements at crossovers and their varying ability to support certain train maneuvers between mainlines, and passenger-train station-platform stop locations.

7.2 Methodology

The methodology for this study investigates the incremental build-out of triple track on what is initially a 240-mile, double-track route between two terminals (Table 7.2). Passenger trains traveling over the route stopped at seven evenly-spaced stations (on either outside track, in the case a third mainline was already in place), with a three-minute station dwell at each.

Table 7.2: Initial route and train characteristics

Route Characteristic	Value
Length	240mi
Crossover Spacing	16mi
Total Number of Crossovers	14
Traffic Control System	2-block, 3-aspect CTC
Train Characteristic	**Value**
Freight Train Consist	115 loads, 6,325ft
Maximum Freight Speed	50mph (40mph through turnout)
Passenger Train Consist	7 coaches, 500ft
Maximum Passenger Speed	110mph (40mph through turnout)
Traffic Composition	heterogeneous, variable

As mentioned above, triple-track installation involves the added complexity of deciding on a particular turnout arrangement at crossovers to provide a train dispatcher with the required train routing flexibility between each mainline track. The two crossover arrangements analyzed in this study are the "parallel" arrangement and the "herringbone" (Figure 7.1).

There are unique advantages, and disadvantages, to each of the two types of crossover arrangements. The *parallel* crossover arrangement allows for simultaneous, parallel train moves between Mainlines 1 and 2, and Mainlines 2 and 3 at a single crossover location. However, there is no way for a train to get all the way from one outside track to the other outside track (i.e.
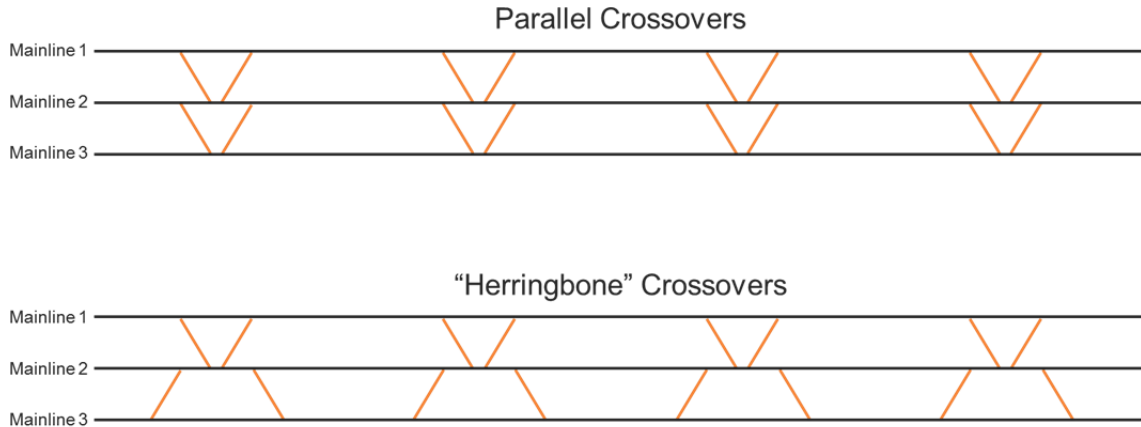
**Figure 7.1: Crossover arrangements on triple track**

Mainline 1 to 3) at a single crossover location with this turnout arrangement. Two crossover moves at successive crossover locations are required to move between outside tracks. This constrains routing flexibility and may be unsuitable for particular corridors where such maneuvers are frequently required. However, the advantage of the *parallel* arrangement is that it allows two parallel moves (Mainline 1 to 2, and 2 to 3) to occur simultaneously at one location; one movement does not interfere with the other, regardless of the direction each train is moving. This allows the middle track to be effectively used as a series of center sidings between the two outside tracks. Trains heading towards each other on the center track in adjacent triple-track sections can simultaneously diverge to either of the outside tracks at a single crossover location without conflicting with the opposing center-track movement.

The *herringbone* arrangement allows for a full train movement from Mainline 1 to 3 (and 3 to 1) at a single crossover location. However, this type of movement requires full occupancy of the crossover, thereby locking out any other simultaneous movements until the train has fully cleared it. This can adversely affect the utility of the crossover and its effective capacity. This arrangement also limits the utility of the center track as a center siding. Opposing train movements on the center track will directly conflict with each other when they converge on a

single crossover location. Regardless of which track the first train is diverging to, the opposing train must stop and wait for the other train to clear the crossover before it can diverge to an outside track. The *herringbone* arrangement is functionally equivalent to the "A-Frame" arrangement also used throughout North American railways.

In this study, these two types of crossover arrangements are presented as alternatives for comparative purposes; however, in practice actual triple track arrangements such as the BNSF route between Chicago and Aurora, Union Pacific in Nebraska, and CSX in Virginia often consist of a combination of *parallel* and *herringbone* crossovers depending on local operational needs. Stations with a single outside platform or connections to branch lines and yards may dictate *herringbone* crossovers to facilitate required moves between outside tracks at a single location and avoid extended running against the predominate current of traffic. On routes where the outside tracks are not signaled for operation in both directions and only the center track is bi-directional, the full-crossover move may not even be a consideration. In these instances, it is more important to provide the capability for *parallel* movements at a single crossover location to increase the utility of the third track. Similarly, crossovers adjacent to stations with island platforms may often require simultaneous crossover moves, favoring *parallel* crossovers for these locations.

In order to isolate the effects of triple-track installation on train delay, a balanced approach was taken to expansion of track infrastructure (Figure 7.2). Triple track was added along the route in an evenly-distributed fashion, such that the route was always laterally balanced with segments of third mainline. This balancing strategy aims to avoid confounded results that may arise from a disproportionate allocation of triple track along particular sections of the route.
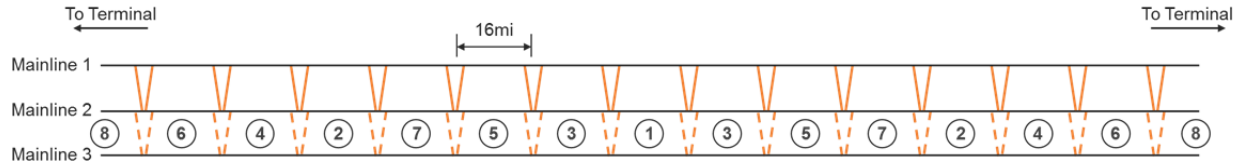
Figure 7.2: Triple track construction order (circled number)

To further clarify, Figure 7.2 shows that, initially, one 16-mile triple-track segment is constructed in the middle of the route. The second step then involves the simultaneous construction of two new triple-track segments, such that the route now has three evenly-distributed sections of third-mainline track. Once these three segments are constructed, each subsequent step includes the paired construction of two new segments of triple track, extending off either the middle segment, or the two outer segments.

7.3 Results

Simulations were run at each level of triple-track installation for three different traffic mixtures on the shared corridor: 48 freight (FRT) + 16 passenger (PAX) trains, 60 FRT + 12 PAX, and 52 FRT + 12 PAX. The results do not include delay data at 100 percent triple track (when each yard is connected through to the third mainline), as the final terminal-to-mainline triple-track connection creates inconsistent simulated delay data that are not representative of delay reductions that would otherwise be experienced. Consequently, capacity benefits associated with triple-track installation are analyzed up until final connections to the two terminals are made.

Delay data generated by the simulations were graphed, while holding traffic mixture and crossover arrangement anonymous (Figure 7.3). The data suggests a linear relationship between the delay per 100 train-miles and the percentage of triple track installed along the route, regardless of either traffic mixture or crossover arrangement.

79
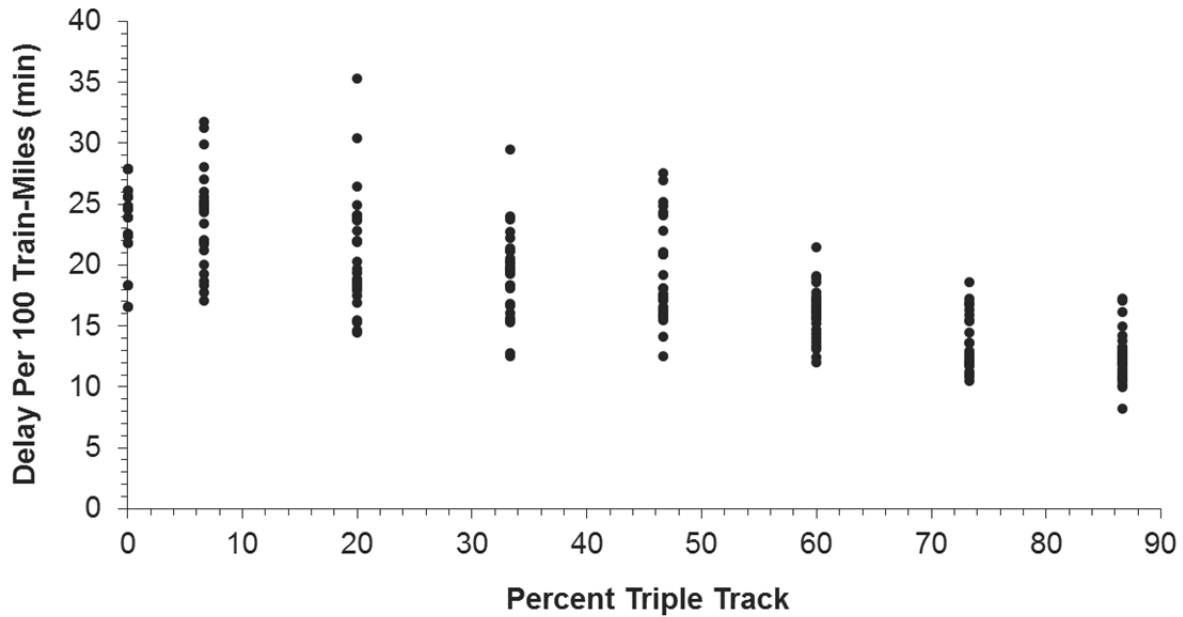
Figure 7.3: Delay as a function of percent triple track

In addition to the linear relationship with negative slope between percent triple track and delay, there is some indication of a reduction in delay variability. Additions of triple track above 60 percent indicate a narrower distribution of delay data, compared to the wider bands at lower percentages of triple track. This suggests uncertainty in delay performance when only a few sections of triple-track are in place. For example, when three triple-tracked segments have been constructed (e.g. 20 percent triple track), the utility of these new segments will depend on train schedule, since fewer triple-track segments means that meets and passes are more likely to occur elsewhere on the route. This effect implies future work in observing the delay response associated with different triple-track build-out strategies; e.g. a continuous addition of triple track from one end of the route to the other or grouped in the center of the route, rather than the balanced distribution used here.

The delay results for each of the two traffic mixtures and crossover arrangements were then considered individually (Figure 7.4). Linear trend lines based on regressed data points
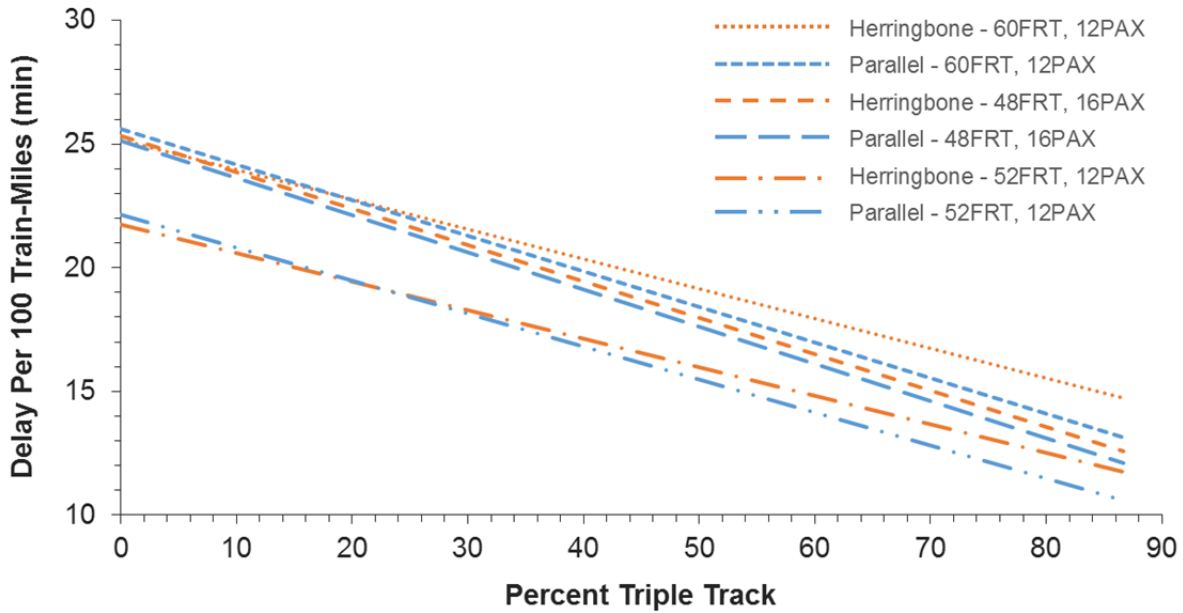
Figure 7.4: Delay as a function of percent triple track, with respect to crossover arrangement and traffic mixture

characterize the triple-track delay response for each combination of crossover and traffic characteristics. The trends shown here are similar to the linearity observed in the single-to-double track research described in Chapter 5.

It is evident that all six trend lines are similar, but there are subtle qualitative patterns. For each traffic condition the *herringbone* crossover arrangement appeared to have slightly higher delay than the *parallel* arrangement. However, given the small difference and lack of statistical tests, no conclusions can be drawn. Furthermore, the small number of conditions tested precludes any generalizations about this relationship except to suggest that further research is needed.

The results are consistent with those previously observed in the transition from single to double track in which delay increases with traffic volume. However, the effect of train volume is less clear when the mix of freight and passenger trains is considered. In particular, the trend lines in Figure 7.4 suggest a similar delay response among the three traffic mixtures, each with

different levels of train heterogeneity that may mask a potentially greater difference between these traffic levels under constant ratios of freight to passenger trains. The bunching of data for the 60FRT-12PAX and 48FRT-16PAX traffic volumes indicates the relatively high influence of adding/subtracting higher-speed passenger operations on the shared corridor, in contrast to the lessened effect of adding/removing freight trains on the line. This notion is directly supported through observation of the relatively isolated 52FRT-12PAX data, which has the same total train count as the 48FRT-16PAX case (64 trains per day). The difference between the two cases is an exchange of four freight for four passenger trains. The reduced delay of the 52FRT-12PAX case supports the notion that passenger train operation disproportionately creates train delay in comparison to freight traffic. The relative influence of additional high-priority, higher-speed passenger trains can be rationalized given that same-speed homogeneous freight traffic could, theoretically, run uninterrupted bi-directionally with just two mainlines. These results emphasize the importance of heterogeneity on train delay – a condition whose significance was studied by Dingler et al. (2013), and is amplified here by the stark speed difference between the freight and passenger trains. Again, these results are only visually suggested, and additional simulation conditions and statistical analysis would be necessary to arrive at any definitive conclusions.

Overall, the results suggest that, irrespective of crossover configuration or traffic mixture studied, triple track reduced delay per 100 train-miles by roughly half (~12min). Over the entire 240-mile route this corresponds to an approximately 30-minute reduction.

7.4 Conclusions

Increasing congestion on North American freight rail corridors, coupled with simultaneous interest in increasing passenger train speed and frequency, suggest the need to better understand the line-capacity benefits of triple-track installation under North American

railroad operating conditions. The results of the experiment design in this preliminary investigation suggest a linear relationship between train delay and percent triple track installed, regardless of crossover arrangement or traffic mixture studied. The results also suggested a possible benefit from a *parallel* crossover arrangement (compared to the *herringbone* arrangement). Overall a 90 percent triple-track installation resulted in a roughly 50 percent reduction in normalized train delay relative to an initial double-track route.

While essential to the North American railroad landscape, research into the incremental capacity in transitioning from double to triple track finds its application within European networks as well. Although most lines in Europe are double track and already support frequent, higher-speed passenger service, there is a desire for operation of longer freight trains over longer distances on this same infrastructure. As the efficiencies of these longer trains are realized, and barriers to international freight interoperability are removed, the European freight rail market share will increase. An increasing number of freight trains on the double-track, passenger-oriented corridors in Europe poses the same concerns regarding the capacity of the existing infrastructure to support future traffic. Under this scenario, European infrastructure owners may face the same prospect of making investments in sections of three-mainline track to incrementally increase line capacity.

CHAPTER 8
CONCLUSIONS AND FUTURE WORK

North American railroads anticipate continued growth in freight traffic and expanded passenger service on freight corridors. In order to avoid congestion with its associated loss in service quality and increased operating costs, railroads need to invest in new and expanded infrastructure. While there are a number of metrics used to measure railway capacity (Sogin 2013), the principal one used in this thesis is normalized train delay. Three major areas of interest were addressed, beginning with an analysis of long-train operations on routes with short sidings, then an analysis of the incremental capacity of transitioning from single to double track, and concluding with an initial consideration of the effect of incrementally adding a third track.

Results from the long-train short-siding analyses concluded that train replacement ratio (i.e. the ratio in the length of long trains to short trains on a route) strongly affects the infrastructure investment required to support operation of long trains while maintaining baseline levels of service. A declining linear relationship exists between train replacement ratio and required investment in siding extensions (i.e. sidings made long enough to accommodate longer trains). For example, routes with a replacement ratio of 3:2 required that roughly half the sidings on a route be extended in order to maintain existing levels of service. A ratio of 2:1 indicated that only about a third of sidings need to be lengthened. The merits of uni-directional, long-train operations were also evident, since this technique showed no adverse effects on train delay, while simultaneously minimizing infrastructure investment.

Where traffic density and mixtures dictate the need for double tracking, project alternatives are often identified using simple practitioner heuristics regarding siding connection length, position, and order. Results from simulation analyses concluded that no one heuristic was definitive; rather, each played a role in affecting train delay. In particular, double-tracking

projects made in the latter half of a full progression from single to double track decreased train-delay more substantially. Longer double-tracking projects showed more consistent delay reduction, while shorter projects showed increased sensitivity to the effects of position and order.

In transitioning from double to triple track, the results suggested a linear relationship between train delay and percent triple track installed, and indicated a slight benefit in the implementation of a *parallel* crossover scheme as opposed to the *herringbone* arrangement. Triple-tracking 90 percent of a double-track route resulted in a roughly 50 percent reduction in normalized train delay.

Although railroads must consider many factors in selecting capital expansion projects, the analyses and guidelines presented here can streamline the decision process by helping to quickly identify projects with the most potential for more detailed engineering evaluation.

While the results presented here shed light on the link between track arrangement and capacity, track construction is a relatively costly alternative to capacity expansion. Consideration should be given to efficient scheduling that maximizes the utility of existing and planned track infrastructure. For example, a siding offers little benefit if trains do not normally meet or pass at its location based on their typical operating schedules. A study that quantifies the interrelationships of train scheduling, track usage, and train delay would be beneficial to a more sophisticated understanding of capacity investment.

Each of the research topics addressed here are also, in one form or another, linked to yard and terminal operations. For example, short yard tracks undermine some of the efficiencies of long-train operation. Additionally, yard operations are a source of train delay regardless of mainline capacity. Integrated modeling of the capacity interaction between yards, terminals, and mainlines should be a high priority in future studies of rail capacity.

REFERENCES

Association of American Railroads. 2008. *Overview of America's Freight Railroads*. URL https://www.aar.org/BackgroundPapers/Overview%20of%20Americas%20Freight%20Railroads.pdf

Association of American Railroads. 2012. *Railroad Facts, 2012 Edition*. Washington D.C.

Atanassov, I., C.T. Dick & C.P.L. Barkan. 2014. Siding spacing and the incremental capacity in the transition from single to double track. In: *Proceedings of the 2014 Joint Rail Conference*, Colorado Springs, USA.

Atanassov, I. & C.T. Dick. 2015a. Capacity of single-track railway lines with short sidings to support operation of long freight trains. Accepted: *Transportation Research Record: Journal of the Transportation Research Board*.

Atanassov, I. & C.T. Dick. 2015b. Delay and required infrastructure investment to operate long freight trains on single-track railway with short sidings. In: *Proceedings of the International Heavy Haul Association Conference*, Perth, Western Australia.

Atanassov, I. & C.T. Dick. 2015c. Influence of siding connection length, position, and order on the incremental capacity of transitioning from single to double track. In: *Proceedings of the 94th Annual Meeting of the Transportation Research Board of the National Academies*, Washington, D.C.

Atanassov, I. & C.T. Dick. 2015d. Incremental Capacity in transitioning from double to triple track on shared rail corridors. In: *Proceedings of the International Association of Railway Operations Research (IAROR) 6th International Seminar on Railway Operations Modelling and Analysis*, Tokyo, Japan.

Barrington, B. & D. Peltz. 2009. 10,000-ft distributed power intermodal trains. In: *Proceedings of the Ninth International Heavy Haul Association Conference*, Shanghai, China, pp. 673-681.

Barton, R. & T. McWha. 2012. *Reducing Emissions in the Rail Sector: Technology and Infrastructure Scan and Analysis*. National Research Council of Canada.

Cambridge Systematics. 2007. *National Rail Freight Infrastructure Capacity and Investment Study*. Prepared for the Association of American Railroads, Washington D.C.

Dick, C.T. & A.M Clayton. 2001. Impact of new railway technology on grain transportation in Western Canada. In: *Transportation Research Record: Journal of the Transportation Research Board*, No.1742, Transportation Research Board of the National Academies, Washington, D.C., pp. 45-53.

Dick, C.T. & L.E. Brown. 2014. Design of bulk railway terminals for the shale oil and gas industry. In: *Proceedings of the 2014 ASCE Shale Energy Engineering Conference*, Pittsburgh, PA.

Dick, C.T. & J.R. Dirnberger. 2014. Advancing the science of yard design and operations with the CSX hump yard simulation system. In: *Proceedings of the 2014 Joint Rail Conference*, Colorado Springs, USA.

Dingler, M., A. Koenig, S. Sogin & C.P.L. Barkan. 2010. Determining the causes of train delay. In: *Proceedings of the 2010 Annual AREMA Conference*, Orlando, FL.

Dingler, M.H., Y-C. Lai & C.P.L. Barkan. 2012. Mitigating train-type heterogeneity on a single-track line. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*. 227(2): 140-147.

Dingler, M.H., Y-C. Lai & C.P.L Barkan. 2013. Effect of train-type heterogeneity on single-track heavy haul railway line capacity. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 228(8): 845-856.

Gussow, J. & N. Welch. 1986. Expansion of Canadian National Railway's line capacity. *Interfaces*. 16(1): 51-64.

Jaumard, B., T.H. Le, H. Tian, A. Akgunduz & P. Finnie. 2013. An enhanced optimization model for scheduling freight trains. In: *Proceedings of the 2013 Joint Rail Conference*, Knoxville, USA.

Kraft, K. 1982. Jam capacity of single track railway lines. In: *Proceedings of the Twenty Third Annual Meeting of the Transportation Research Forum*, New Orleans, USA. 23(1): 461-471.

Krueger, H. 1999. Parametric modeling in rail capacity planning. In: *Proceedings of the 1999 Winter Simulation Conference*, pp. 1194–1200.

Lai, Y-C. & C.P.L Barkan. 2011. Comprehensive decision support framework for strategic railway capacity planning. *Journal of Transportation Engineering*, 137(10): 738-749.

Lindfeldt, O. 2006. Influences of station length and inter-station distance on delays and delay propagation on single-track lines with regional rail traffic. *Computers in Railways X: Computer System Design and Operation in the Railway and Other Transit Systems*, WIT Press, pp. 511-520.

Lindfeldt, O. 2010. *Railway Operation Analysis: Evaluation of Quality, Infrastructure, and Timetable on Single and Double-Track Lines with Analytical Models and Simulation*. KTH Royal Institute of Technology, Department of Transport, Stockholm, Sweden. (Doctoral Dissertation).

Lindfeldt, O. 2012. 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*, WIT Press, pp. 313-324.

Martland, C.D. 2008. Railroad train delay and network reliability, *AAR Research Report R-991*, Transportation Technology Center, Inc., Pueblo, Colorado.

Martland, C.D. 2013. Introduction of heavy axle loads by the North American rail industry. *Journal of the Transportation Research Forum*, 52(2): 103-125.

Mitra, S., S. Mitra & D. Tolliver. 2010. Estimation of railroad capacity using parametric methods. *Journal of the Transportation Research Forum*, 49(2): 111-126.

Moore Ede, B., A. Polivka & J. Tunna. 2007. Optimizing capacity by trade-offs among train control, infrastructure, mechanical, and engineering considerations. In: *Proceedings of the 2007 IHHA Specialist Technical Session*, Kiruna, Sweden, pp. 413-421.

Newman, R.R., A.M. Zarembski & R.R. Resor. 1991. Economic implications of heavy axle loads on equipment design operations and maintenance. In: *Proceedings of the American Society of Mechanical Engineering*, WAM, RTD-Volume 4, Rail Transportation.

Shih, M-C., C.T. Dick, S.L. Sogin & C.P.L. Barkan. 2014. Comparison of capacity expansion strategies for single-track railway lines with sparse sidings. *Transportation Research Record: Journal of the Transportation Research Board*. No.2448, Transportation Research Board of the National Academies, Washington, D.C., pp. 53-61.

Sogin, S. 2013. *Simulations of Mixed Use Rail Corridors: How Infrastructure Affects Interactions Among Train Types*. University of Illinois at Urbana-Champaign, Urbana, USA. (Master's Thesis).

Sogin, S., Y.C. Lai, C.T. Dick & C.P.L. Barkan, 2013a. Analyzing the incremental transition from single to double track railway lines, In: *Proceedings of the International Association of Railway Operations Research (IAROR) 5th International Seminar on Railway Operations Modelling and Analysis*, Copenhagen, Denmark.

Sogin, S., Y-C. Lai, C.T. Dick & C.P.L. Barkan. 2013b. Comparison of capacity of single- and double-track rail lines. *Transportation Research Record: Journal of the Transportation Research Board*. No.2374, Transportation Research Board of the National Academies, Washington, D.C., pp. 111-118.

Tobias, A., D. House & R. Wade, 2010. Planning capacity improvements in the Chicago-Milwaukee-Madison rail corridor using the Rail Traffic Controller (RTC) rail operations simulation model, In: *Proceedings of the 2010 ASME/IEEE/ASCE Joint Rail Conference*, Urbana, USA.

Wilson, E. 2015. *Rail Traffic Controller (RTC)*, Berkeley Simulation Software, LLC.