CHAPTER 8
RAIL TRANSIT CAPACITY

1. User's Guide
2. Mode and Service Concepts
3. Operations Concepts
4. Quality of Service Concepts
5. Quality of Service Methods
6. Bus Transit Capacity
7. Demand-Responsive Transit
8. **Rail Transit Capacity**
9. Ferry Transit Capacity
10. Station Capacity
11. Glossary and Symbols
12. Index

CONTENTS

LIST OF EXHIBITS

1. INTRODUCTION

Many factors contribute to the number of trains that can be operated and the number of people that can be carried over a given time period on a rail transit line or railroad corridor—the fundamental determinant of the capacity of the line. These factors are related to vehicles, station characteristics, signaling system technology, and operational characteristics.

Organization of Chapter 8.

Chapter 8 of the *Transit Capacity and Quality of Service Manual* (TCQSM) presents methods for calculating the capacity of a variety of rail modes and right-of-way types.

- Section 2 introduces the fundamental concepts and factors associated with rail capacity.
- Section 3 describes the basic operation of train control and signaling systems and their relationship to the minimum train headway.
- Section 4 describes operational measures and platform design aspects that can improve train capacity, speed, and reliability.
- Section 5 provides the computational methods for calculating rail capacity for various modes and configurations of rail transit systems, including methods for measuring or estimating input values.
- Section 6 presents potential applications of this chapter's methods and describes the role of simulation in rail capacity analysis.
- Section 7 provides examples of applying this chapter's computational methods.
- Section 8 is a list of references used to develop the material in this chapter.
- Appendix A provides substitute exhibits in metric units for Chapter 8 exhibits that use U.S. customary units only.

The majority of the material in this chapter first appeared in *TCRP Report 13: Rail Transit Capacity* (*1*). Although written in the mid-1990s, this report remains the leading reference on the topic. The basic principles of rail capacity have not changed, although improvements to rail vehicles and technology have occurred and continue to occur, and the methodologies presented here are flexible enough to accommodate these changes.

HOW TO USE THIS CHAPTER

At a minimum, be familiar with the content of Sections 2 and 3 before applying this chapter's computational methods.

Sections 2–4 of this chapter build upon the general transit capacity, speed, and reliability concepts presented in Chapter 3, Operations Concepts, providing information specific to rail transit. Readers will ideally be familiar with the contents of Sections 2 and 3 before trying to apply this chapter's computational methods. Section 4 provides information about rail system design and operations that are useful to consider when planning a potential new rail system (e.g., as part of an alternatives analysis).

Section 5 begins with a general methodology for estimating the capacity of a rail transit line in terms of both trains and persons per hour. Although some of the equations may look complicated, the calculations are straightforward substitutions of input values for each variable in the equation. The majority of the effort is in selecting appropriate values to apply to the equations; this section provides guidance in this area, but assumes familiarity with the basic rail capacity concepts from Sections 2 and 3.

Refinements to the general methodology are subsequently presented for commuter rail and automated guideway transit (AGT) lines, and a separate method is presented for estimating the capacity of ropeway modes (e.g., aerial trams and funiculars). Section 7 provides examples of performing the computations associated with these methods.

Section 6 describes potential applications of this chapter's methods to planning applications. When greater detail is required, or the operations of a rail line are more complex (e.g., lines merging or crossing, mixed freight and passenger operations), simulation is typically used to evaluate operations and determine maximum reliable train throughput. Section 6 includes sections on the role of simulation and provides examples of its application to rail capacity analysis.

OTHER RESOURCES

Other TCQSM material related to rail transit capacity includes:

- The "What's New" section of Chapter 1, User's Guide, which describes the changes made in this chapter from the 2nd Edition.
- The "Rail Transit" subsection of Chapter 2, Mode and Service Concepts, which defines and describes the various rail submodes.
- Chapter 3, Operations Concepts, which presents general capacity concepts applicable to all transit modes, including rail.
- The "Passenger Load" subsection of Chapter 5, which presents a detailed method of estimating railcar passenger capacity applicable to any railcar dimensions, seating arrangement, and transit agency loading policy. The length-based method presented in Chapter 8 assumes generic light rail and heavy rail car dimensions and relatively comfortable standing passenger loads, which may not be applicable to specific situations.
- Chapter 10, Station Capacity, which presents methods for sizing station platforms and their exits. Crowded platforms can slow down passenger boarding and alighting, which increases dwell times and potentially reduces a rail line's capacity.
- The manual's CD-ROM, which includes spreadsheets for applying the general rail capacity method and for estimating the capacity of single-track bi-directional operation. It also includes links to electronic versions of all the TCRP reports referenced in this chapter.

The CD-ROM accompanying the manual includes spreadsheets for evaluating rail transit capacity.

2. RAIL CAPACITY FUNDAMENTALS

OVERVIEW

Ideally, station dwell time and the minimum train separation produced by the signaling system will control line capacity, but other factors may need to be considered.

Rail transit systems encompass a variety of technologies, vehicle sizes, and applications. Despite these variations, a few basic factors—in particular, dwell times and train signal control systems—typically control the number of trains that can be operated along a section of a line during an hour. The number of cars per train and the diversity of passenger demand control how many people those trains can carry.

LINE CAPACITY

Line capacity and vehicle capacity, both relating to the number of trains that can be operated per hour, are equivalent terms for rail capacity.

Line capacity is the maximum number of trains that can be operated over a section of track in a given period of time, typically 1 h. Ideally, the combination of the train signaling system being used and the station with the longest dwell time will control the line capacity. However, under less than ideal conditions, any of a number of other factors may control line capacity. These include:

- Signaling systems designed for the minimum planned train headway, rather than the maximum capacity practical;
- Speed restrictions due to sharp curves or steep downgrades on the approach to the station with the longest dwell time;
- Line crossings and merges, particularly at-grade track junctions;
- Time required to turn back a train at a terminal station; and
- Mode-specific issues, such as light rail trains operating in mixed traffic or commuter rail trains sharing tracks with freight trains.

Propulsion power constraints can also limit line capacity.

The factor providing the lowest capacity—the weakest link—will constrain the capacity of a given section of a line.

Train Control and Signaling

Streetcars and portions of some light rail systems that operate at low speeds do not use train signals. Multiple trains may be allowed to berth at stations where space permits.

Most rail modes rely on signaling systems to maintain safe separation between trains. The minimum distance between trains must be long enough for a train to come to a complete stop, with a suitable safety margin between it and the train ahead. All urban rail transit train control systems are based on dividing the track into sections known as *blocks* and ensuring that trains are separated by a suitable and safe number of blocks. The longer the time required for a train to traverse (pass through) a block—whether due to long block lengths, low train speeds, or station dwell time—the longer the minimum headway between trains, and the lower the line capacity. Train control is discussed in detail in Section 3. Exhibit 8-1 illustrates the operation of a typical three-aspect (red/yellow/green) signal system.

Exhibit 8-1
Basic Train Signal Operation

Note: R = red, Y = yellow, G = green.

Dwell Time

Dwell time at the station with the highest passenger volumes often will control line capacity.

Dwell time is frequently the dominant factor in determining the minimum train headway and, thus, the line capacity. The three main components of dwell time are:

- Door open and close time, and time waiting to depart once the doors close;
- Passenger flow time; and
- Time the doors remain open after passenger flow ceases.

Of these three factors, passenger flow time is the largest and the hardest to control. It depends on passenger volumes at stations, the number of doors on a train, the door widths, the level of crowding inside the train and on the platform, and congestion between boarding and alighting passengers at the train door. The other two factors are, to a great degree, under an agency's control. Minimizing the time spent in a station without passenger flows occurring is important in maintaining reliable train operations, particularly when a line is operating near capacity.

Operating Margin

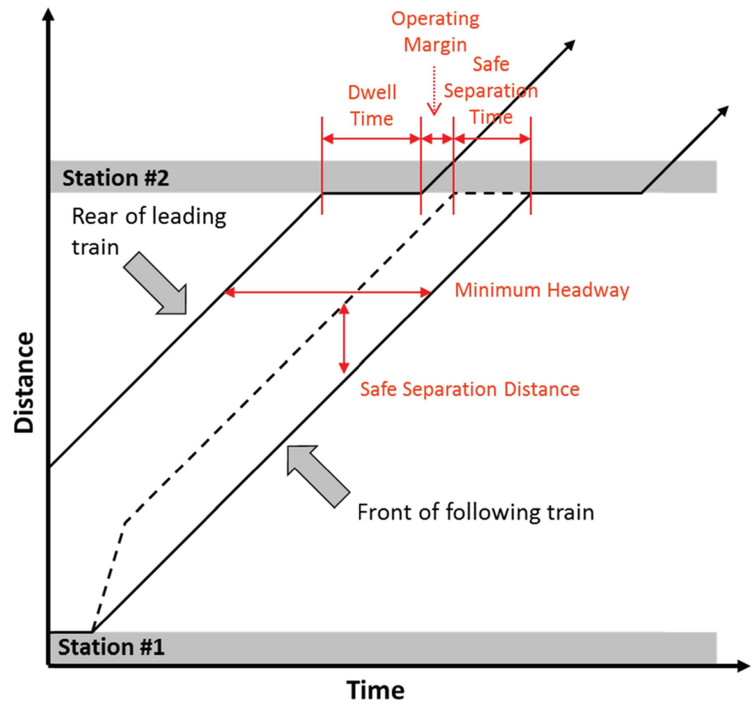
An operating margin is "slack time" built into the minimum headway to accommodate small irregularities in service. If a train is late by more than the operating margin, following trains will be delayed.

When a rail system is operating close to its capacity, small irregularities in service can lead to delays, as a train is not able to approach a station until the train ahead departs. These irregularities can be caused by variations in station dwell times, variations in train performance, and—on manually driven systems—variations between operators. To compensate for these variations, when creating a minimum headway, most rail systems add an operating margin to the combination of the signal system's minimum train separation time and the critical station dwell time. The operating margin is, in effect, the amount of time a train can run behind schedule without interfering with the following trains and, consequently, is an important component of line capacity. Operating margins are discussed further in Section 5.

Non-interference Headway

In most cases, the combination of the safe separation time imposed by the train control and signaling system, the longest (or *critical*) average dwell time along the line, and the operating margin will determine the minimum headway that can be operated along the route. This minimum headway is known as the *non-interference headway,* because as long as it can be maintained (i.e., actual dwells do not exceed the average dwell plus the operating margin), following trains will be able to proceed from one station to the next without stopping or slowing for preceding trains, as shown in Exhibit 8-2.

Exhibit 8-2
Distance–Time Plot of Two Consecutive Trains



Source: *TCRP Report 13* (*1*).
Note: Acceleration and braking curves omitted for clarity.

If a train's dwell exceeds the average dwell plus operating margin, however, the following train will need to slow or stop to maintain the required safe separation distance and will not be able to approach the next station at its planned speed. This delay, in turn, will force the next train to slow or stop to maintain its required separation, creating a cascade of delays to following trains that will be extremely difficult to resolve as long as trains continue to arrive at the minimum headway.

As can be seen, train operation at the minimum headway can be easily disrupted. Transit agencies that operate rail lines at or near the minimum headway therefore try to manage station dwell times—for example, through the use of timers visible to train operators, and through passenger education efforts to encourage passengers to step aside to allow others to exit the train first and to not hold train doors open.

Guideway Characteristics

Turnbacks

A typical terminal station will have a center (island) platform, allowing passengers to board trains on either side. A number of designs are possible, but a common, lower-cost (but also potentially capacity-constraining) design is to locate a crossover in advance of the station. This crossover allows entering trains to be sent to either platform, and exiting trains to be sent to the correct departure track. When a line operates at short headways, the amount of time required to load and unload passengers, and for the operator to change ends, inspect the train, and check train integrity and braking will be longer than the headway between trains. As a result, a second train will arrive and occupy the other platform while the first train is still preparing to depart. A capacity constraint will result if the first train is unable to clear the crossover before a third train arrives to use the platform that the first train is vacating. Exhibit 8-3 shows this process.

Alternative terminal station designs are discussed in Section 4.

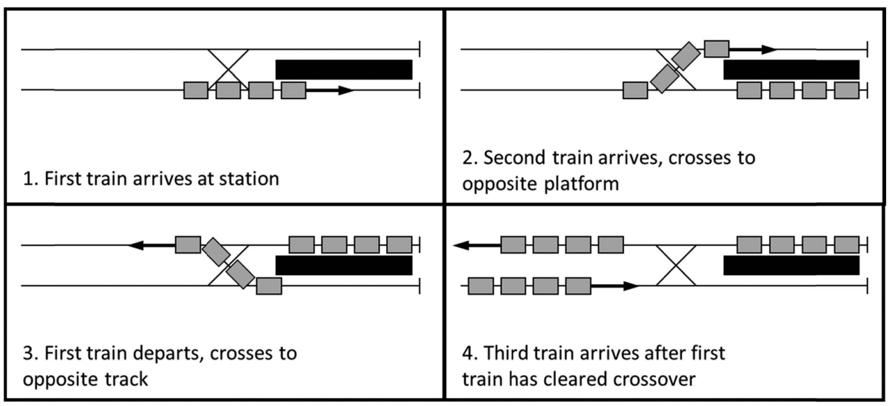


Exhibit 8-3
Turnback Operation with Crossover Located in Advance of Station

Alternative configurations also are possible, including far-side crossovers and tail tracks beyond the terminal station, turning loops, and turning pockets (a third track in between the two mainline tracks, for turning selected trains at a point before the end of the line).

As described in Section 5, when turnbacks are correctly designed and operated, they should not control capacity on a new rail system. However, turnbacks can be a constraint on older systems, where physical constraints—particularly in subways—may have resulted in less-than-optimal designs, or when passenger demand has generated the need for more service than the system was originally designed for.

Junctions

Locations where lines merge, diverge, or cross at-grade can constrain capacity, or introduce the likelihood of interference, when scheduled headways approach 2 to 2.5 min. Two trains may need to use the space where the tracks cross, but only one train can occupy that space at a time. The minimum interval between trains on a given line at an at-grade ("flat") junction is a combination of:

1. The time required for an opposing train to move through the junction,
2. The time required to move ("throw") and lock the switches,
3. The delay incurred in decelerating from and accelerating to line speed, and
4. The minimum headway imposed by the signaling system on the line.

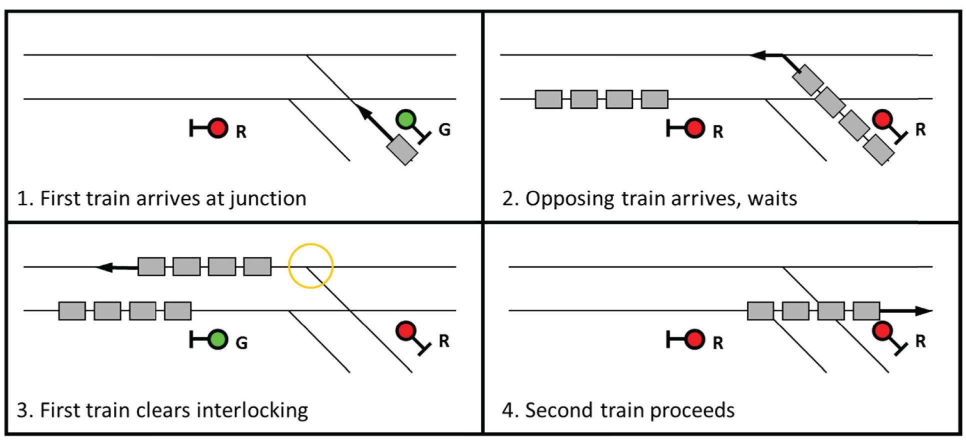
Conceptually, the process is similar to that used for calculating headway based on dwell time at a station, in that both headways are based on the minimum train separation on the lines plus the time a train is stopped. In this case, time stopped is spent waiting for another train at a junction rather than waiting to serve passengers.

It is not desirable for one train to have to wait for another. When more capacity is required, grade-separated ("flying") junctions are typically used. Exhibit 8-4 depicts the two types of junctions. Exhibit 8-5 illustrates the operation of a flat junction. Section 5 discusses junctions in more detail.

Exhibit 8-4
Types of Rail Junctions

(a) Flat (Pittsburgh) (b) Flying (Paris)

Exhibit 8-5
At-Grade ("Flat") Junction Operation



Note: R = red, G = green.

Vehicle–Platform Interface

The performance of trains while stopped at stations has a significant effect on overall line capacity and can, in many cases, be the controlling element. Factors affecting dwell times include:

- The volumes of passengers boarding and alighting from trains during peak hours;

- The physical configuration of the platform: its width, length, curvature, usable area for passenger queuing and circulation, and configuration and capacity of vertical circulation;
- The rate at which passenger alight from and board the train;
- The extent of any horizontal gaps between train door sill and the platform edge or differences in elevation between the platform and the car floor—which affects the rate at which passengers board and alight;
- The time required to open and close the train doors; and
- Operational procedures affecting the boarding process.

Ideally, platforms should have tangent (straight) edges, with the surface of the platform at the same level as the train vestibule, to meet the requirements of the Americans with Disabilities Act (ADA) for level boarding with no more than a three inch gap between train door and platform edge. Platforms also should extend the full length of the train, so that doors on all cars open onto the platform. This configuration optimizes the flow capacity of the train doors and minimizes the length of time required to unload and board a train's passenger loads. Platforms should be wide enough to allow boarding passengers to queue on the platform while allowing adequate lanes for alighting passengers to exit the train and walk to the vertical circulation elements. Stairs, escalators, ramps, and/or elevators should be provided in sufficient numbers and spaced along the platform to allow the platform to be cleared of arriving passengers prior to the arrival of the next train. However, level boarding is not always possible or practical, especially on existing commuter rail and light rail systems not designed for level boarding.

Vehicle Characteristics

The characteristics of the rolling stock also affect line capacity. Doorway flow rates are a function of the number, size, and spacing of doors on the train and the interior vestibule space available to passengers boarding and alighting. These flow rates, in turn, influence dwell time.

Mode-Specific Issues

The line capacity factors identified above are applicable to most major rail modes, particularly heavy rail (rail rapid transit), and one of these factors will generally control line capacity. Sometimes, though, issues unique to a particular mode may need to be considered as well:

- *Light Rail*—single-track operations, on-street operations (either in mixed traffic or in an exclusive right-of-way), street-level boarding, and the characteristics of traffic signal priority.
- *Commuter Rail*—mixed freight and passenger operations, limits on the number of trains imposed by the owner of the tracks being used, differences in locomotive power, single-track operations, and boarding from low-level platforms.
- *Automated Guideway Transit*—widely varying technology, potential for off-line stations that allow trains to bypass stations and other trains.
- *Ropeway*—line length, line speed, vehicle or carrier spacing.

PERSON CAPACITY

Person capacity defined.

Person capacity is the maximum number of people that can be carried in one direction over a section of track in a given period of time, typically 1 hour, under specified operating conditions without unreasonable delay, hazard, or restriction, and with reasonable certainty.

The definition of person capacity is less absolute than the definition of line capacity, as it depends on the number of trains operated, the length of those trains, passenger loading standards, and variations in passenger demand between trains and between individual cars of a given train.

The theoretical capacity is the number of cars per hour per direction, times the maximum design load of each car.

Person capacity accounts for variations in passenger arrivals and distribution, and is lower than the theoretical capacity.

This last factor, known as *loading diversity,* provides an important distinction between a line's theoretical capacity and a more realistic person capacity that can actually be achieved on a sustained basis. The theoretical capacity assumes that all the offered capacity can be used by passengers. In practice, this only occurs when a constant queue of passengers exists to fill all available seats and standing room—a situation that is undesirable in a transit operation, as it leads to crowded platforms and passenger delay. Transit passengers generally do not arrive at an even rate over the course of an hour, and generally do not distribute themselves evenly among the cars of a train. Accounting for loading diversity allows one to determine the number of people that can be accommodated during an hour without pass-ups occurring.

How many people can be carried vs. how many people could be carried.

Constraints on staff and equipment resources must also be considered. Line capacity considers how many trains *could* be operated, assuming no constraints on the supply of cars to form trains, nor any constraints on the number of operators available to drive those trains. Knowing, and designing for, a line's ultimate person capacity is often important in long-term planning. However, it may be just as important to know in the short term how many trains *can* be operated and the person capacity of those trains, given existing resources.

Passenger Loading Standards

The passenger-carrying capacity of a transit or rail car, for both seated and standing passengers, is a critical element of person capacity. Peak train loads are estimated based on loading standards that are either developed by the operating agency, based on equipment specifications and assumptions about appropriate design loads, or derived from rules of thumb. For rail transit systems, design capacity includes full occupancy of any seats, plus an allowance for a certain number of standees at a reasonably comfortable quality of service. For commuter rail systems, nominal capacity typically assumes full seated capacity without standees. Chapter 5, Quality of Service Methods, provides guidance on determining appropriate design loads.

See Chapter 5 for guidance on passenger loading levels.

Loading standards are typically based on maximum design loads—the maximum number of people that can be accommodated at a specified quality of service. Crush loading represents the physical capacity of the vehicle to accommodate passengers and loads greater than the maximum design load. The former is expected to be a regular everyday occurrence, while the latter may be tolerable for short periods on an infrequent basis when delays occur or when trains are cancelled—as the system recovers to its normal operating state—or when extraordinary holiday or special event loadings occur.

Loading Diversity

Passengers do not load evenly into cars and trains over the peak hour. Three different types of loading diversity have to be considered: (*a*) loading diversity within a car, (*b*) loading diversity among cars of a train, and (*c*) unevenness of passenger demand during the peak hour.

Loading diversity within a car.

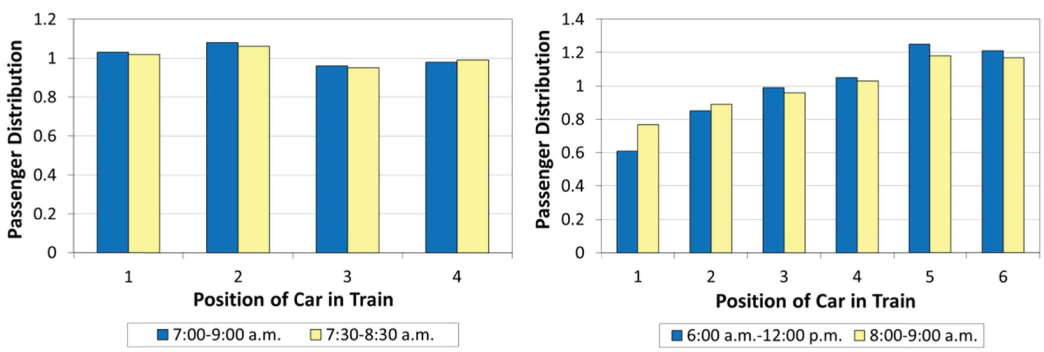
The first type of loading diversity is within a car. In individual cars, the highest standing densities occur around doorways while the lowest densities occur at the ends of the cars. Several European urban rail systems add doors, sometimes only single-stream, at the car ends to reduce this unevenness.

Loading diversity within a train.

A second type of diversity occurs in uneven loading among cars of a train. Cars that are closer to station exits and entrances will be more heavily loaded than more remote cars. This inefficiency can be minimized by staggering platform entrances and exits between ends, centers, and third points of the platforms. This is not always possible or practiced. Even so, relatively even loading often occurs due to the duress factor that encourages passengers to spread themselves along the platform during heavily traveled times—or risk being unable to get on the next arriving train.

Vancouver TransLink (formerly BC Transit) has measured car loadings at a station where passengers are regularly passed up, as shown in Exhibit 8-6(a).

Exhibit 8-6
Illustrative Peak-Hour Passenger Distribution Between Cars of Trains



(a) Vancouver, SkyTrain (Broadway Station) (b) Toronto, Yonge Subway (Wellesley Station)

Source: *TCRP Report 13* (*1*).
Notes: 1.0 represents an individual car load equal to the average load of all cars in the train.
Vancouver data collected inbound direction Oct. 27, 1994, 50 trains, 6,932 passengers.
Toronto data collected southbound direction Jan. 11, 1995, 99 trains, 66,263 passengers.

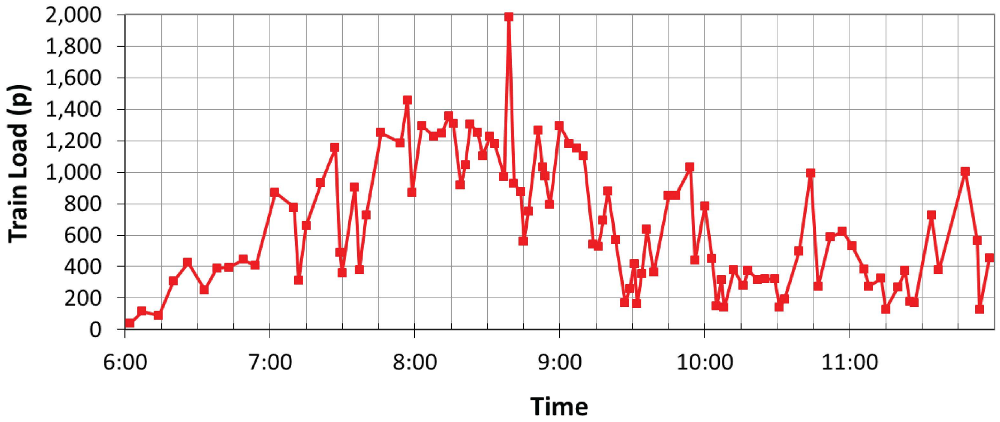
In Vancouver, there is no significant variation in the average loading diversity between cars of a train in either the peak hour or the peak 2-hour period, both of which are within the range of +5% of an average (mean) load to -6%. However, the imbalance between cars on individual trains ranges from +61% to -33%. The average evenness of loading can be attributed to four factors: short trains, wide platforms, close headways, and dispersed entrance/exit locations among the system's stations.

Toronto's Yonge Street subway has a more uneven average loading between cars than Vancouver, as seen in Exhibit 8-6(b). During the morning peak period, the rear of the train is consistently more heavily loaded. This pattern reflects the dominance of the major transfer station at Bloor Street, with the interchange occurring at the rear (northern) end of the Yonge subway platform. As would be expected, there is less variation in the average car loading diversity between the peak hour and the peak morning period due to the pressures on passengers to spread along the platforms at busy times. The average diversity of individual car loading over the peak period has a range of +26% to -39%. The imbalance for cars on individual trains ranges from +156% to -89%.

Loading diversity within the peak hour.

The third and most important type of diversity is the unevenness of passenger demand over the peak hour. Passengers do not arrive evenly and uniformly on any rail transit system, as shown dramatically over the extended peak period in Exhibit 8-7 for Toronto's Yonge Street subway. This exhibit shows the realities of day-to-day rail transit operation. The morning peak 15 min has a pronounced abnormality at 8:35 a.m. following a short gap in service. The different loading, train by train, is significant, and it is difficult to visually pick out the peak hour or the peak 15 min.

Exhibit 8-7
Example A.M. Train Loads, Toronto

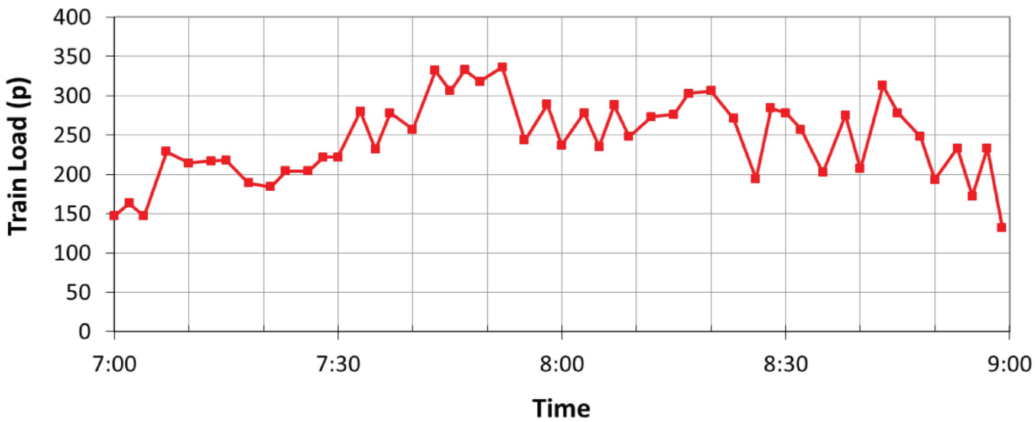


Source: *TCRP Report 13* (*1*).
Note: Data collected January 11, 1995 at Wellesley Station in the southbound direction.

Short-turn services turn back before the end of the line, providing more frequent service in the middle of the line than at the ends.

Exhibit 8-8 shows an a.m. peak period for Vancouver's SkyTrain that, although without major delays, shows the irregular loading from train to train due to the interlace of short-turn trains with regular service from 7:30 a.m. onward.

Exhibit 8-8
Example A.M. Train Loads, Vancouver



Source: *TCRP Report 13* (*1*).
Note: Data collected October 27, 1994 at Broadway Station in the inbound direction.

Number of Cars

The number of cars in a train is a major determinant of person capacity—the longer the train, the more people it can carry. However, there are limits to how many cars can be added to a train, set by the lengths of platforms, the supply of cars, and (for light rail) city block lengths.

Platform Lengths

Station platforms are designed for the longest train the system plans to operate. When platforms are located above or below grade, they are difficult to lengthen once constructed. In some instances, for example, at New York's South Ferry subway station (before it was rebuilt) and some older commuter rail stations, the platform is shorter than the train length, and passengers wishing to exit trains must do so from selected cars only at the front or rear of the train. However, this kind of operation is not generally desirable and is not typical practice for new systems.

Car Supply

Even when the platform design allows for longer trains, a shortage of rail cars may preclude operating longer trains. This kind of constraint is typically financial—new rail cars averaged $1.9 to 3.6 million each in 2010–2011, depending on the rail mode and type of car (*2*); additional staff are also required to maintain added cars.

Maintenance and Yard Storage Facilities

The capacity of facilities where rail cars are maintained and stored when not in service can constrain the number of cars available to form trains. Capacity-enhancement plans that consider lengthening station platforms also need to consider the potential need to expand rail yards and enlarge maintenance facilities. For many existing systems, lengthening yard tracks may be infeasible, difficult, or expensive.

Propulsion Power

For commuter rail systems that employ locomotives, the number of cars that can be accommodated on the train is limited by the horsepower and other performance characteristics of the locomotive. Many traditional locomotives in commuter service are capable of handling trains of up to eight cars. Some newer models of high-horsepower locomotives are capable of pulling 10- to 12-car trains.

For electric traction rail modes, power supply limitations can constrain the number of cars and trains that can use a given track section. Electrical substations are located at intervals along a line, each of which is capable of powering only a certain number of cars within its section of track. Therefore, even though the train control system may provide the capability of operating short headways continually, the electrical system may only support that capability for short periods of time (i.e., until the number of cars that one substation can power in a given track section has been reached).

Street Block Lengths

Street block lengths can be a major limitation for at-grade systems that operate on-street. Most jurisdictions are unwilling to allow stopped trains to block intersections and so require that trains not be longer than the shortest street block where a stop is likely. This issue is especially noteworthy in Portland, where unusually short street blocks of 200 ft (65 m) in the downtown area limit trains to two cars. Sacramento has been an exception to the street block length rule and can operate four-car trains in the peak hours. These long trains block one intersection when stopped. This situation arose as the single-track nature of the original Sacramento line (since addressed by double-tracking) imposed a minimum headway of 15 min on the service. The capacity limitation of this headway restriction was therefore partially made up by the operation of relatively long trains.

Exclusive lanes can mitigate street block length constraints.

Street block length is also an issue when another vehicle occupies the lane used by light rail trains. If a vehicle in the lane would cause the rear of the train to protrude into an intersection, then the train would need to wait for the lane to clear before advancing. This issue provides a strong argument for providing an exclusive light rail transit lane where street running with long trains occurs. Indeed, as a result of this concern, operation with mixed traffic is very rare on new light rail transit systems, although it does occur on modern streetcar lines in the U.S., which typically operate one-car trains. Where buses and light rail transit trains operate alongside each other, such as in downtown Calgary and Portland, the rail stations, bus stops, and lanes are laid out to cause minimum interference between the modes.

DESIGN CAPACITY

The generally accepted methods by which capacity is calculated define the maximum possible throughput of trains on a line as the theoretical capacity. Several factors contribute to the design capacity of a line being a somewhat lower level of throughput. Research on following headways for a high-density rail operation in tunnels approaching New York City (*3*) provides a good overview of the factors that go into the estimation of theoretical and design capacity. The results of this research are summarized below.

Train signaling systems are discussed in detail in Section 3.

For capacity measurement purposes, the theoretical headway at a specific signal location is the shortest time a following train can pass that signal location at the same speed as the first train. It generally assumed that the speed is the maximum speed allowed by the best available signal aspect (or speed code if a cab signaling system is employed). For fixed-block signal systems, both cab and wayside, the theoretical headway at a signal or for a signal block is determined by calculating the time required for a train traveling at line speed to clear the signal block, plus an additional *clearing time,* which is defined to be the time it takes for the signal aspect or cab-signal code to return to its highest-speed signal aspect or code after the first train passes (typically on the order of 3 s), plus an allowance for the reaction time of the engineer or operator of the following train to recognize that the preceding signal has cleared (typically 3–4 s).

The clearing time at a given signal is determined by both constant (fixed) and variable factors. Fixed factors include the block length, design speed through the block, and the worst-case train safe-braking distance. The variable factors in determining signal clearing time include the speed and station stopping pattern of the first train as it passes through the control line of that signal. If there are multiple train types and station-stopping patterns, a clearing time can be calculated for each one. This can either be done in the field with a stopwatch or by using a train performance calculator programmed with the line's civil characteristics (grades, curves, civil speed restrictions, maximum speed allowed by the signal code or aspect, and any underspeed assumptions), the unique acceleration and braking characteristics of each train type, and the dwell times at each scheduled station stop within the span of the control line.

Train capacity at a given signal location is calculated by dividing one hour (3,600 s) by the theoretical signal headway (seconds per train), which includes both signal clearing time and engineer reaction time as noted above, to give a maximum throughput expressed in trains per hour. Where there are multiple train types and stopping patterns, a weighted average headway is calculated based on the proportion of each train type relative to the total number of trains.

In commuter rail, just as the signal with the longest clearing time will define segment line capacity, the slowest moving train in a segment will define line capacity, because it will block trains behind it, preventing the faster trains from clearing at optimal speed unless passing sidings or additional tracks are available to prevent delays to the faster trains. These issues arise in local/zone-express type operating patterns, and also in mixed freight/commuter corridors.

It should be noted that the capacity calculated for each signal location assumes that a following train is arriving at that signal ready to accept the newly cleared code or aspect 3 s after it clears. Where a line segment has multiple signals, the longest clearing time defines the capacity of that segment. While some signals in a given segment may clear faster than others, trains operating on close headways cannot take advantage of these fast-clearing signals because their natural progression will reflect the slowest clearing signal.

The capacity of a given line segment, therefore, is based on the *theoretical line headway,* which is the maximum of the headways calculated (or observed) at each signal on the line—sometimes referred to as the *ruling* or *constraining headway.* Throughput is calculated by dividing the theoretical line headway into one hour in the same manner as for an individual signal block.

Typically, plotting capacity with respect to speed results in a bell-shaped curve in which capacity is low at low speeds due to the long time it takes to traverse fixed blocks, high at speeds in the 45–60 mi/h (75–100 km/h) range when supported by appropriate intermediate-speed signal aspects or codes, and tapering down again at higher speeds, which require longer safe braking distances. Braking distances increase with the square of the speed, resulting in much longer signal blocks at higher speeds.

Normally the longest clearing time for a train making station stops is associated with a signal whose control line extends into a station platform. Station dwell time is usually the factor limiting capacity in such cases.

Where there are multiple train types or stopping patterns, an average theoretical line headway is derived from the longest clearing time for each combination of train type and stopping pattern, with a weighted average calculated based on the proportion of each type. In cases where the longest clearing times are at different signal locations for different stopping patterns, the weighting of clearing times is based on the number of times trains of one unique pattern follow trains of another unique pattern.

In addition to the two major variables of train type and stopping pattern, there are other variable factors that determine the actual clearing time behind each train, namely variations in engineer and equipment performance that affect rates of acceleration, deceleration, and maximum speed. These are not typically included in clearing time calculations, but any measure of capacity must account for them in some way.

Calculated capacity at a given signal that does not factor in these variable elements is often referred to as the calculated or theoretical capacity at that signal location.

Each of the variables not included in theoretical capacity can have a negative effect on capacity in one of two ways:

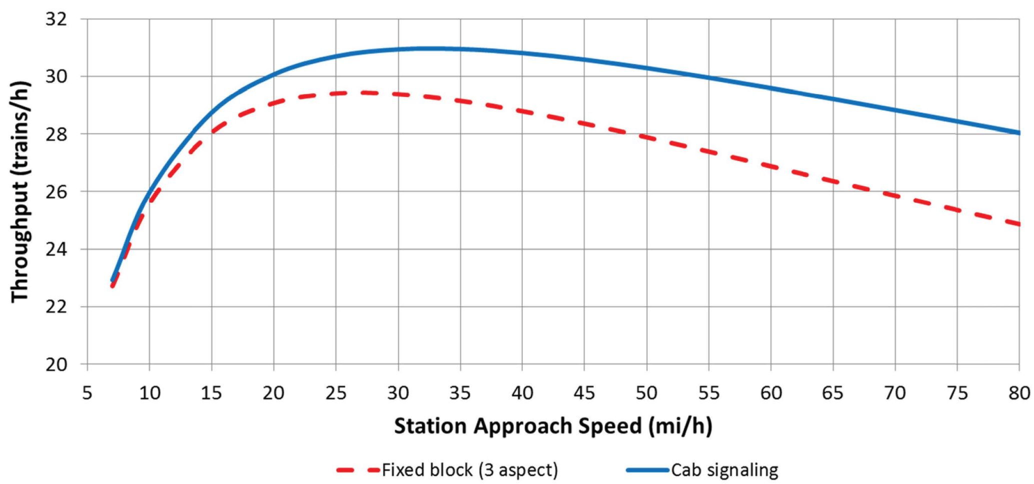
1. By slowing a train as it approaches the ruling signal, making it unavailable to take immediate advantage of the clear signal when it becomes available. In the absence of an operating margin (schedule slack) between trains, this effect will cause every subsequent train that attempts to follow at the theoretical ruling headway to be late by the same amount of time that the first train was delayed.
2. By slowing or delaying a train within the ruling control line such that the signal takes longer to restore to its best aspect or code.

A capacity measure that includes all of the variable factors not included in theoretical capacity can be considered design or practical capacity. For line segments where trains are operating at close to the maximum authorized speed, design capacity generally is estimated to be in the range of 75 to 80 percent of theoretical capacity (i.e., a reduction in theoretical capacity by 20 to 25 percent). This is a rule of thumb that is commonly applied to rail systems for purposes of developing train schedules, projecting future growth requirements, and performing capacity analyses.

SPEED

The capacity of a transit line in terms of train throughput is affected by the speed at which the transit vehicles travel. As speed increases, the distance and time required to safely stop a train also increases, which means that trains need to be spaced further apart in order to operate safely. This relationship between speed and throughput capacity is illustrated in Exhibit 8-9. Cab signal systems and communications-based train control systems are able to more efficiently track the position of trains on a line and therefore offer slightly more line capacity than traditional fixed-block signal systems. The actual relationship between speed and throughput capacity will be a function of the specific design characteristics of the signaling system, so the graph should be considered as illustrative.

Exhibit 8-9
Illustrative Capacity as a Function of Speed for a Rail Transit Line



Source: Calculated with this chapter's methods.
Note: Assumes 45-s average dwell time and 20-s operating margin.

For rail transit systems, the speed at which capacity is maximized is in the range of 25 to 30 mi/h (40 to 50 km/h, usually significantly lower than the allowable top speed), and the peak throughput usually is in the range of 30 trains per hour (equivalent to an approximate 2-min headway). At speeds higher than this, maximum speed and throughput are inversely proportional to each other. Operating at a lower speed for the sake of higher throughput may be acceptable for a heavy rail or light rail transit line in an urban environment but these speeds generally are slower than desired for commuter rail or express transit operations. As a result, these latter systems may intentionally be operated at a lower design capacity than could otherwise be achieved, in order to achieve safe operations at higher speeds.

Some transit systems (e.g., PATH) achieve throughput greater than 30 trains per hour.

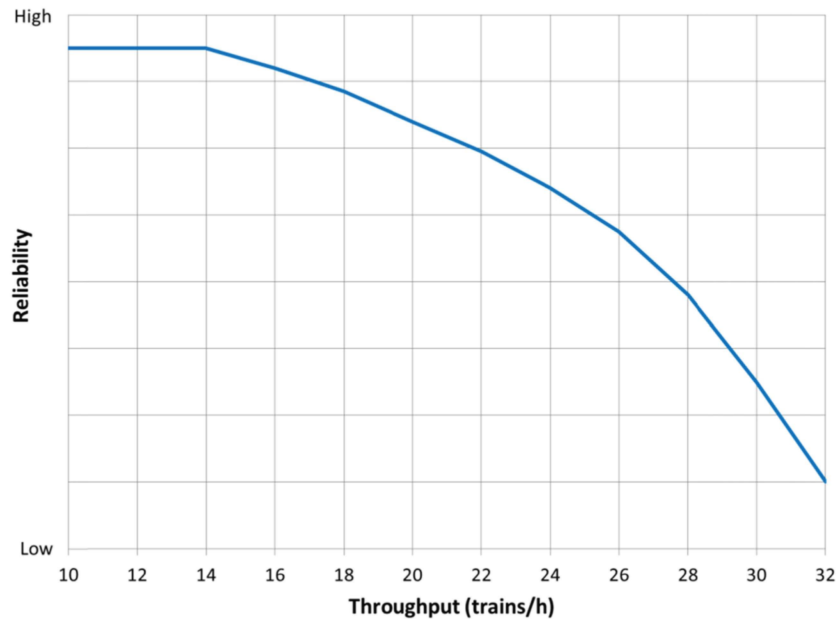
POSITIVE TRAIN CONTROL

Commuter rail systems, which operate on the national railroad network, are subject to federal regulations for positive train control (PTC), a system that improves the safe operation of trains by overriding the discretion of the train engineer to pass a stop signal or operate at a higher-than-permissible speed. At the time of writing, the regulations and associated system design concepts continue to be under development. The effect of such systems on capacity is unclear. Well-designed systems should be able to implement PTC without a significant degradation of capacity, particularly for passenger-only systems or systems with relatively few types of trains in operation. Complex systems with mixed freight and passenger operations and multiple types of trains may see a decrease in capacity, since the PTC system would need to protect against reasonable worst-case operating conditions. The costs, difficulties, and impacts of retrofitting PTC to existing systems also are unclear, but will need to be part of the analysis for projects that change the traffic mix or seek to increase the capacity of an existing system. Simulation models, described in Section 6, provide a tool by which the capacity impacts associated with implementing PTC can be determined and factored into a capacity analysis.

RELIABILITY

The design capacity of a rail line is also related to the desired level of operational reliability. If the frequency of service is relatively low, individual trains have time to recover from minor delays without impacting other trains on the line. As a line approaches its capacity, however, it becomes vulnerable to the condition in which a delay to a single train causes additional delays that cascade or propagate to other trains. The magnitude of the total delay and the time required to recover to normal operating conditions increases as the density of traffic increases, as shown in Exhibit 8-10. There are several ways in which operational reliability can be measured, and a quantitative relationship between reliability and design capacity is difficult to measure, so the reliability axis on the chart is not dimensioned. The shape of the curve, however, is generally considered to be roughly as shown.

Exhibit 8-10
Conceptual Reliability–Throughput Relationship



Transit and rail lines that are asked to deliver a level of throughput close to their theoretical capacity can expect to be less reliable than those that are planned to operate at longer headways. Systems for which predictable, highly reliable service is paramount should be planned for a lower design capacity to better ensure the line's ability to recover from individual train delays.

3. TRAIN CONTROL AND SIGNALING

OVERVIEW

Functions of signaling.

The role of signaling is to safely separate trains from each other and protect specific paths through interlockings (switches) at junctions and crossovers. Additional functions include automatic train stops, should a train run through a stop signal, and speed control to protect approaches to junctions, sharp curves, and approaches to terminal stations where tracks end at a solid wall.

Early forms of surface rail transit, such as streetcars, operated without signal systems—on a *line-of-sight* basis—where the operator of the transit vehicle was responsible for maintaining a safe distance from the preceding vehicle. Some existing systems that operate at relatively low speeds continue to operate on a line-of-sight basis, which can deliver a relatively high vehicle throughput with the trade-off of a top speed lower than rail transit operations equipped with signaling systems.

Signaling technology is very conservative, but signaling cannot protect from every eventuality.

Rail transit signaling maintains high levels of safety based on emergency brake applications and fail-safe principles ensuring that no single failure—and often multiple failures—should allow an unsafe event. The rigor with which fail-safe principles have been applied to rail transit has resulted in an exceptional safety record. However, the safety principles do not protect against all possibilities, including possible human error. An increasing inability to control the human element—responsible for three-quarters of rail transit accidents or incidents—has resulted in new train control systems using automation to reduce or remove the possibility of human error.

Automatic train control adds further features to the train protection of basic signaling, including automatic driving and train supervision that regulates service.

Automatic train control.

This section describes and compares the separation capabilities of the following types of rail transit train control systems: fixed block, cab, and moving block. It is applicable to the main rail transit grouping of electrically propelled, multiple-unit, grade-separated systems.

FIXED-BLOCK SYSTEMS

Track circuits.

In a fixed-block system, trains are detected by the wheels and axles of a train shorting a low-voltage current inserted into the rails. The rails are electrically divided into blocks. The blocks will be short where trains must be close together (e.g., in a station approach), and can be longer between stations where trains operate at speed.

Fixed-block systems provide a coarse indication of train location.

The signaling system only knows the position of a train by the simple measure of block occupancy. It does not know the position of the train within the block; it may have only a fraction of the train, front or rear, within the block. At block boundaries, the train will occupy two blocks simultaneously for a short time.

A minimum of two empty blocks is required between trains for a two-aspect fixed-block system.

A three-aspect system (red/yellow/green) is typical; four-aspect systems (red/yellow/flashing yellow/green) are also found on some commuter rail systems.

In the simplest two-aspect (red/green) block system, the signals display only stop (red) or go (green). A minimum of two empty blocks must separate trains and these blocks must be long enough for the braking distance plus a safety distance. The simplest system can accommodate a throughput approaching 24 trains per hour. This does not provide sufficient capacity for some high-volume rail lines. Higher capacity can be obtained from combinations of additional signal aspects, shorter block lengths, and overlay systems that electronically divide blocks into shorter "phantom" sections—for trains equipped for this overlay.

Conventional train control systems can support a throughput of 30 trains/track/hour.

In this way, conventional train control systems can support a throughput of up to 30 trains per hour with typical train lengths, performance, controlling station dwells, and operating margins. Overlay systems can increase this throughput by 10 to 15%. A notable exception to this is in Russia where conventional signaling routinely handles 40 metro trains per hour. This is achieved by tightly controlling station dwells to a maximum of 25 s and rigorous adherence to schedule using digital clocks in each station to display the seconds from the departure of the previous train. Newer Moscow metro lines have been designed for 44 and 48 trains per hour—by far the closest train spacing on any rail system, irrespective of technology.

Requiring a driver to control a train's speed and commence braking according to multiple-aspect color light signaling requires considerable precision to maximize throughput. Cab signaling provides assistance in this regard and reduces capital and maintenance costs.

CAB SIGNALING

Cab signaling sets authorized, safe train speeds.

Cab signaling uses electronic codes inserted into each track circuit and detected by an antenna on each train. The code specifies the maximum allowable speed for the block occupied and may be termed the *reference* or *authorized* speed. This speed is displayed in the driver's cab—often so that the authorized speed and actual speed can be seen together.

The authorized speed can change while a train is in a block, as the train ahead proceeds, allowing drivers to adjust train speed close to the optimum with less concern about overrunning a trip stop. Problems with signal visibility on curves and in inclement weather are reduced or eliminated. Cab signaling avoids much of the capital and maintenance costs of multiple-aspect color light signals, although it is prudent and usual to leave signals at interlockings and occasionally on the final approach to and exit from each station.

Reducing the number of color light signals makes it economically feasible to increase the number of aspects and it is typical, although not universal, to have the equivalent of five aspects on a cab signaling system. A typical selection of reference speeds would be 50, 40, 30, 20, and 0 mi/h (80, 70, 50, 35, and 0 km/h).

MOVING-BLOCK SYSTEMS

Moving-block signaling systems are also called *transmission-based* or *communication-based* signaling systems. A moving-block signaling system can be compared to a fixed-block system with very small blocks and a large number of aspects. However, a moving-block signaling system has neither blocks nor aspects. The system is based on continuously or frequently calculating the clear (safe) distance ahead of each train and then relaying the appropriate speed, braking, or acceleration rate to each train.

Moving-block signaling is based on the use of target points.

This system requires continuous or frequent two-way communication with each train, and precise knowledge of a train's location, speed, and length, and of fixed details of the line—curves, grades, interlockings, and stations. With this information, a computer can calculate the next stopping point of each train—the *target point*—and can command the train to brake, accelerate, or coast accordingly. The target point will be based on the normal braking distance for that train plus a safety distance.

The safety distance is the maximum distance a train can travel after it has failed to act on a brake command before automatic override (or overspeed) systems implement emergency braking.

Without track circuits to determine block occupancy, a moving-block signaling system must have an independent method to accurately locate the position of the front of a train, and then use look-up tables to calculate its end position from the length associated with that particular train's identification. The first moving-block systems used a wire laid alongside or between the running rails, periodically transposed from side to side. The wire transmitted signals to and from antennas on the train, while counting the transpositions determined location.

The use of exposed wayside wires is a maintenance problem and refinements use inert transponders located periodically along the track. These are interrogated by a radio signal from each train and return a discrete location code. Positioning between transponders relies on the use of a tachometer. Communications to and from the train are then radio-based, with protocols to ensure safety and reliability and that messages are received by, and only by, the train they are intended for.

Communication can be made secure.

The computers that control a moving-block signaling system can be located on each train, at a central control office, dispersed along the wayside, or a combination of these. The most common arrangement is a combination of onboard and central control office locations.

Safety Issues

Safety on rail transit is a primary consideration when rail systems are designed. It encompasses all aspects of design, maintenance, and operations. In fixed-block signaling, electrical interlockings, switch, and signal setting are controlled by relay logic. A rigorous discipline has been built around this long-established technology, for which processor-based controls are now finding a role.

A moving-block signaling system is inherently processor-controlled. Processor-based train control systems intrinsically cannot meet the fail-safe conventions of traditional signaling. Computers, microprocessors, and solid-state components have multiple failure opportunities and cannot be analyzed and tested in the same way as conventional equipment. Instead, an equivalent level of safety is provided based on statistical failure modes of the equipment. Failure analysis is not an exact science. Although not all failure modes can be determined, the statistical probability of an unsafe event can be predicted.

HYBRID SYSTEMS

There are times when an urban rail transit system shares tracks with other services, such as long-distance passenger trains, whose equipment is impractical or uneconomic to equip with the moving-block signaling system. Hybrid or overlay systems are available that allow use by unequipped trains—with longer separation—while still obtaining the close headway of the moving-block system for the urban or short-distance trains.

Hybrid systems can allow equipment not equipped for moving block operation to operate on lines signaled with moving blocks.

AUTOMATIC TRAIN OPERATION

Automatic acceleration has long been a feature of rail transit, where relays, and more recently microprocessors, control the rate of acceleration smoothly from the initial start to maximum speed. Linking this feature to onboard commands from the signaling system provides automatic train operation.

Automated train operation systems often also provide for manual operation.

The driver or attendant's role is typically limited to closing the doors, pressing a train start button, and observing the line ahead, with limited manual operating capabilities to deal with certain failures. Dispensing entirely with a driver or attendant was controversial when introduced but has demonstrated its economy and safety on numerous automated guideway transit (AGT) systems, and on selected rail systems in Europe and Canada.

Automated train operation may provide a 2 to 4% capacity increase.

Automatic train operation (ATO), with or without attendants or drivers, allows a train to follow the optimum speed envelope more closely and commence braking for the final station approach at the last possible moment. This reduces station-to-station travel times, and, more importantly, from the point of capacity, it minimizes the critical station *close-in time*—the time from when one train starts to leave a station until the following train is berthed in that station. This can increase total line capacity by 2 to 4 percent.

AUTOMATIC TRAIN SUPERVISION

Automatic train supervision (ATS) is generally not a safety-critical aspect of the train control system. At its simplest, it does little more than display the location of trains on a mimic board or video screen in the central control or dispatcher's office. Increasing levels of functionality are available.

Corrective measures to correct late running trains.

In more advanced systems where there is ATO, computer algorithms are used to attempt to automatically correct lateness. These are rare in North America and are generally associated with the newer moving-block signaling systems.

Predictive control.

A further level of ATS strategies is possible: predictive control, when a computer looks ahead to possible conflicts (for example, a merge of two branches at a junction). The computer can then adjust terminal departures, dwell times, and train performance to ensure that trains merge evenly without holds.

The non-vital ATS system can also be the host for other features such as onboard system diagnostics and the control of station and onboard information through visual and audio messages, including those required by the Americans with Disabilities Act (ADA).

ON-STREET PREFERENTIAL TREATMENTS

Rail preferential treatments on city streets constitute a more limited set of treatments than treatments to accommodate buses, given the constraints posed by a dedicated trackway for streetcars and at-grade light rail transit. This limits the total freedom of vehicle movement that trains can achieve, to take advantage of bypass maneuvers and sudden shifts in vehicle routing. In addition, light rail trains are longer than buses and can have a greater impact on signal operations and the ability to provide effective transit signal priority. Given the higher cost of streetcar and at-grade light rail transit operation, greater emphasis is placed on providing a dedicated right-of-way within a street section if possible. This includes a predominance for light rail transit for median transitways, which can be provided separate from adjacent general traffic, as opposed to shared-use lanes with general traffic. While streetcars can also operate in separate transitways, most streetcar operations share lanes with general traffic and tend to institute transit signal priority in a similar manner as buses due to their shorter train length (typically only one vehicle).

Transitways

A variety of light rail and streetcar transitway types can be placed within a street right-of-way, providing different levels of protection from other street users. These transitway types can be broadly divided into three groups: (*a*) exclusive, (*b*) semi-exclusive, and (*c*) non-exclusive. Allowable train speeds are higher and potential conflicts with other roadway users are fewer at higher levels of exclusivity (*4*).

An *exclusive transitway* provides complete separation of transit vehicles and other street users. The transitway can be located completely above or below grade, or can run mostly at-grade (with barriers or fences protecting the transitway) with under- or overpasses used at major cross streets.

A *semi-exclusive transitway* allows access across the transitway at controlled locations. Examples of this type of transitway include light rail tracks in the street median, in a protected right-of-way along one side of the street, and along transit malls. Traffic signals are used at locations where cross-street traffic is allowed, and pedestrians and bicycles cross at designated locations only.

A *non-exclusive transitway* allows uncontrolled access onto or across the transitway. Examples include light rail tracks operating in transit-only lanes that can be crossed by turning vehicles at driveways and unsignalized intersections, and mixed-traffic operation (typical streetcar operation).

Examples of these transitway types are provided in the Operating Environments section of Chapter 2, Mode and Service Concepts.

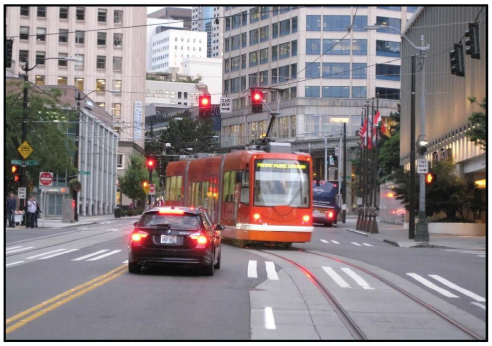
Traffic Signal Priority for Transit Vehicles

Traffic signal priority for LRT or streetcars can be applied under certain signal and traffic operating conditions, more applicable to isolated signalized intersections or where the signal timing in a closely spaced set of intersections (such as in downtown areas) can provide a window for some green extension for transit. Given the typical longer train length for LRT compared to individual bus lengths, a signal preemption strategy may be considered as an alternative to signal priority (including extension of green time), particularly at isolated signalized intersections. Streetcars are shorter than LRT trains—often with just one car per train in U.S. applications—and therefore have the potential to employ signal priority strategies similar to those for buses.

Special traffic signal phases for trains (Exhibit 8-11[a]) can be inserted when LRT vehicles or streetcars turn or queue jump in front of parallel motor vehicle traffic. In many of these cases, the train or streetcar will be operating on the right side of street (whether in a side-of-road alignment or in mixed traffic, and because of the greater turning radius requirement of the vehicles, greater green time associated with the special phase will be required.

When traffic signal priority is applied, stations located near intersections are preferably located on the far (departure) side of the intersection, such that trains can take advantage of the green extension or added signal phase to proceed through the intersection before servicing the station. In these cases, there should preferably be at least one lane available for parallel general traffic, so that motorized vehicles do not queue up behind a stopped train and block the intersection.

Exhibit 8-11
On-Street Rail Preferential Treatment Examples



(a) Queue jump (Seattle)

(b) Curb extension (Little Rock)

Curb Extensions/Boarding Islands

If LRT vehicles or streetcars are running on the curb side of the street, either curb extensions (Exhibit 8-11[b]) or boarding islands can be applied as pedestrian refuge areas to serve as stations. Boarding islands are typically provided where there is a right turn lane and hence the need for a raised pedestrian refuge area between the right turn lane and adjacent through lane where the train or streetcar is operating. In either configuration, the length and width of the station should be adequately sized to accommodate the design train and design passenger demand.

4. TRAIN OPERATIONS

OVERVIEW

The previous section focused on the effects of train control and signaling on the capacity of a rail transit line. This section discusses additional operating issues affecting capacity.

Uniformity of train performance.

There is considerable uniformity of performance of the electrical multiple-unit trains that handle more than 90% of all U.S. and Canadian rail transit, and this uniformity can be further assisted by the widespread introduction of electronic controls and automatic driving. However, there still can be up to a 10% difference in performance between otherwise identical trains due to manufacturing tolerances, aging of components, and variances in set-up parameters, and—particularly on manually driven systems—due to variations in driving techniques between drivers.

Operating margins.

To accommodate these routine irregularities, two allowances are made in rail transit operations planning and scheduling. An *operating margin* is added to the minimum train separation time and maximum load point station dwell time to create a minimum headway. This operating margin is, in effect, the amount of time a train can run behind schedule without interfering with the following trains. The operating margin is an important component in determining the design capacity.

Schedule recovery.

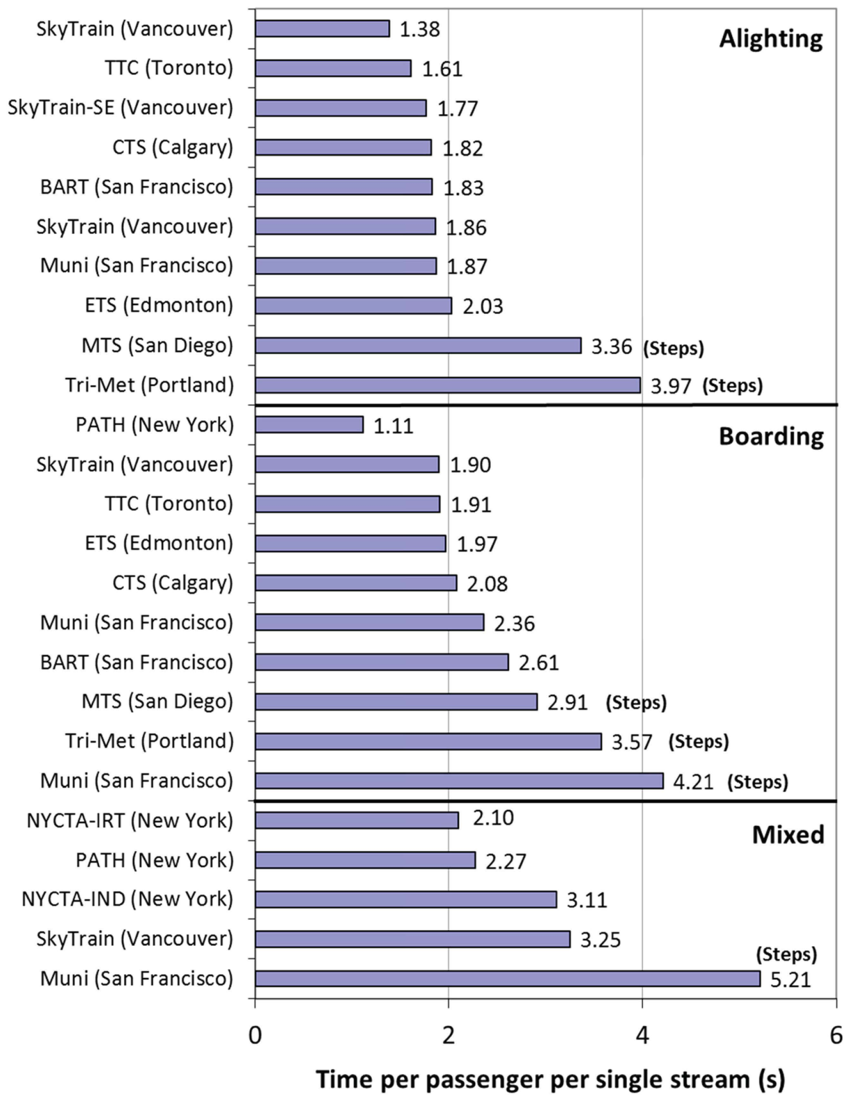
The second allowance is *schedule recovery,* an amount of time added to the terminal turn-around time to allow for recovery from accumulated delays on the preceding trip. Schedule recovery time has some effect on capacity and also has economic implications, as it can increase the number of trains and staff required to transport a given volume of passengers.

DOORWAY FLOW RATES

Flow time is the time in seconds for a single entering or exiting passenger to cross the threshold of the rail transit car doorway, per single stream of doorway width. Extensive rail transit door flow rate data collection took place in 1995 as part of *TCRP Report 13: Rail Transit Capacity* (*1*). Data were collected from a representative set of high-use systems and categorized by the type of entry—level entry being the most common, followed by light rail with door stairwells, with and without fare collection at the entrance. The data sets were partitioned into mainly boarding, mainly alighting, and mixed flows. This work was supplemented in the early 2000s with data collection specific to low-floor light rail cars as part of the development of the TCQSM 2nd Edition (*5*). The results are summarized in Exhibit 8-12.

Exhibit 8-12
Selection of Rail Transit Door Flow Times (1995)

Data in this exhibit are illustrative of the range of values that are possible. Analyses of actual systems should utilize recent empirical data from the lines being analyzed, where available.

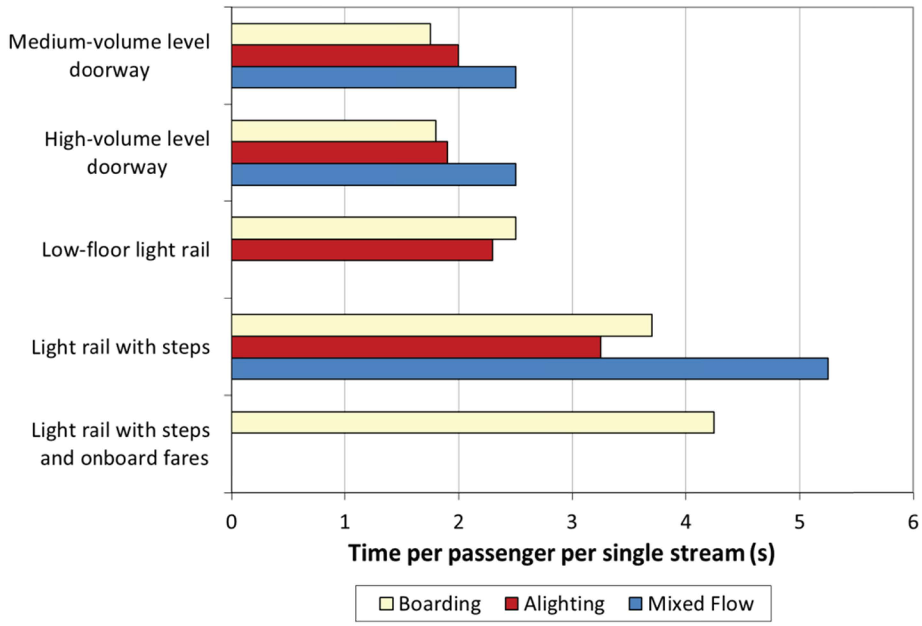


Source: *TCRP Report 13* (*1*).
Note: Level boarding, except up or down steps where indicated.

Passengers ascend steps into a light rail vehicle faster than they descend them on exit.

An interesting result is that passengers enter high-floor light rail vehicles faster from street level than they exit. The overall fastest flow rate, 1.11 seconds per passenger per single stream, was observed on PATH when passengers were boarding empty trains at the Journal Square station in Newark in the morning peak. These flow data are consolidated and summarized by type of flow in Exhibit 8-13.

Exhibit 8-13
Summary of Rail Transit Average Door Flow Times



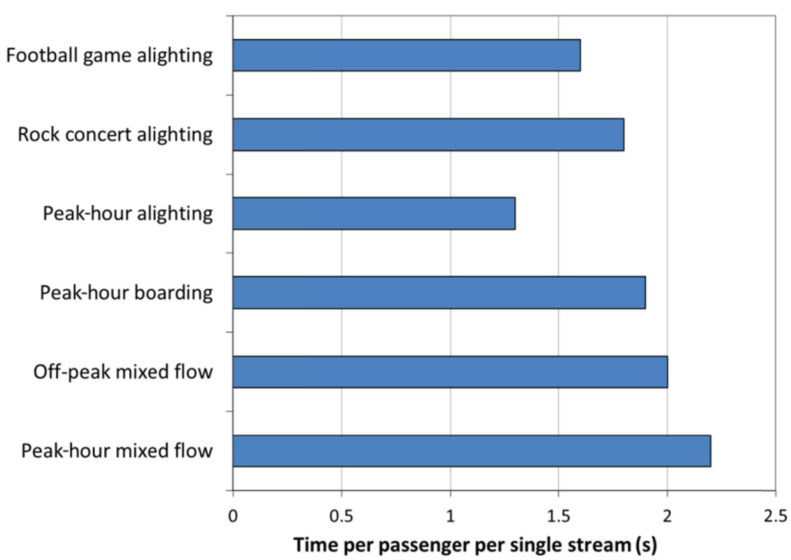
Sources: *TCRP Report 13* (*1*) and TCQSM 2nd Edition (*5*).

The results show that, in these averages, there is little difference between the high-volume systems —older East Coast heavy rail transit—and the medium-volume systems—newer light rail and heavy rail transit. Doorway steps approximately double times for all three categories: mixed flow, boarding, and alighting. Manual fare collection adds about 1 to 3 s per passenger transaction (for specific values, refer to Exhibit 6-4 in Chapter 6, Bus Transit Capacity, page 6–7).

Doorway steps double boarding and alighting times.

While most of the field data collection on doorway flow rates was done during peak periods, off-peak and special event flows were observed on Vancouver's SkyTrain and compared with peak-period flows, as summarized in Exhibit 8-14.

Exhibit 8-14
SkyTrain (Vancouver) Door Flow Rate Comparisons



Source: *TCRP Report 13* (*1*).

Special event passenger rates were found to be slower than weekday peak rates.

Special event flows were observed before a football game, before a rock concert, and on a busy suburban station in the early afternoon base period. The resultant data are contrary to the supposition that special event crowds move faster and that off-peak flows are slower than in the peak hour. BC Transit (now TransLink) has also measured car occupancy differences between normal peak hour operation and after service delays. In the ensuing pressure to travel after a delay, passenger space dropped almost by half from a mean of 3.8 ft^2 per passenger to 2.2 ft^2 per passenger (2.8 p/m^2 to 5 p/m^2).

OPERATING MARGINS

Examples of North American Operating Margins

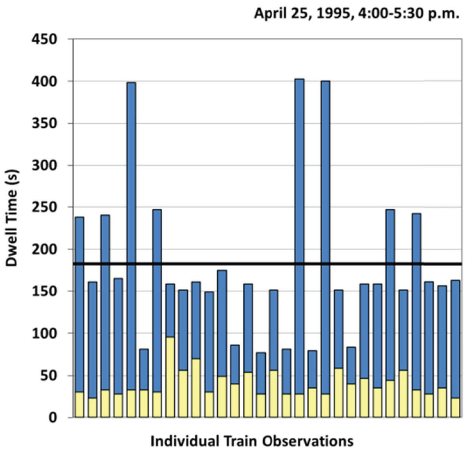
As a starting point for recommending a suitable operating margin, the operating margins incorporated into the schedules of existing systems can be reviewed. The maximum load point, peak period, station dwell time, and headways for several rail transit lines are presented in Exhibit 8-15.

Operating margin examples.

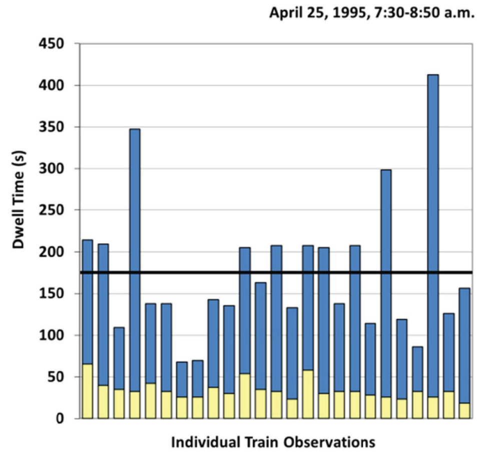
The headways in Exhibit 8-15 for Calgary are all multiples of the 80-s traffic signal cycle. The seemingly erratic headways in Calgary are misleading as three routes, forming two interlaced services, shared this downtown bus and light rail mall. The exhibit also shows the dwell and headway regularity of interlaced services on Vancouver's fully automatic SkyTrain. San Francisco's BART is also automatic, but allows drivers to control when the doors close; it shows a relatively regular headway pattern, but more irregular dwell times.

The final two charts in Exhibit 8-15 show the range of dwell and headway irregularities on manually driven systems. These are not typical of most heavy rail lines throughout the day, but represent lines at or near capacity at the peak point in the peak period. It is at these times that operating margin and schedule recovery times are most needed to correct service irregularities.

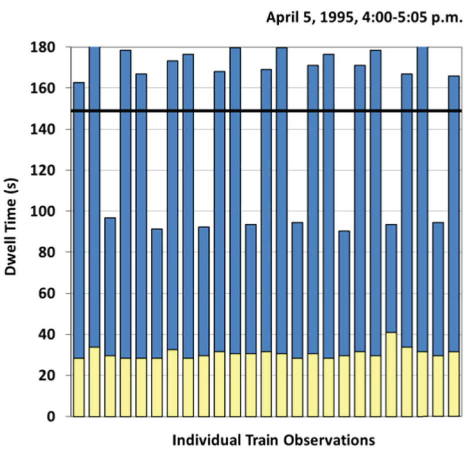
Exhibit 8-16 shows the headway components with the final column indicating the residual time that is a surrogate for the operating margin.



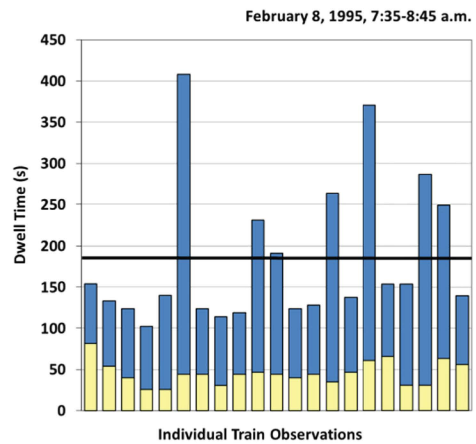
(a) CTS 3rd St. SW EB (Calgary)



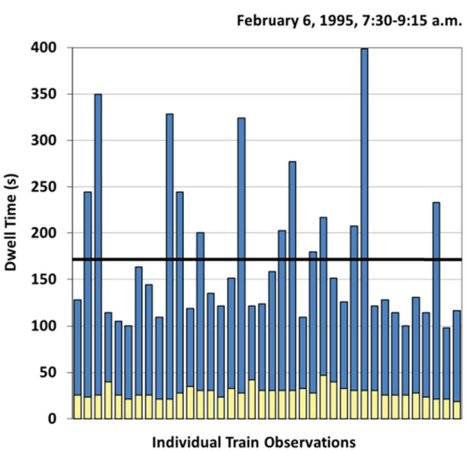
(b) CTS 1st St. SW WB (Calgary)



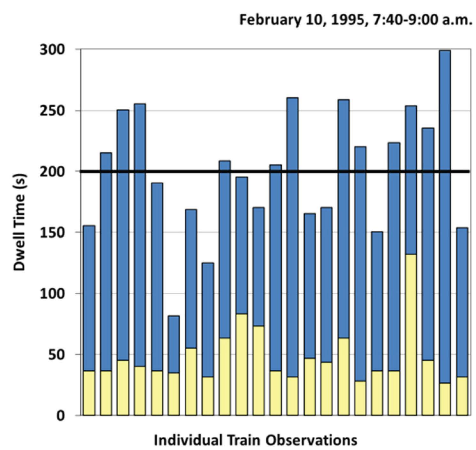
(c) SkyTrain Broadway EB (Vancouver)



(d) BART Embarcadero WB (San Francisco)



(e) TTC King SB (Toronto)



(f) PATH Journal Square WB (Newark)

Source: *TCRP Report 13* (*1*).
Note: EB = eastbound, WB = westbound, SB = southbound.

Exhibit 8-15
Observed Rail Headways and Dwell Times

Dwell Time (s)
Headway (s)
Avg. headway (s)

Light rail headways on observed systems were generally sufficiently long that any irregularities reflected problems other than schedule interference between trains. One of the closest on-street headways is in Calgary, shown at the top. Note that the scales of the graphs vary.

Additional examples of these dwell/headway charts are contained in Chapter 6 of TCRP Report 13 (*1*).

Exhibit 8-16
Dwell and Headway Data Summary of Surveyed Rail Transit Lines Operating at or Close to Capacity (1995)

System and City	Station and Direction	Average Station Dwell (s)	Dwell Standard Deviation (s)	Average Headway (s)	Dwell as % of Headway	Train Control Separation (s)	Estimated Operating Margin (s)
BART San Francisco	Embarcadero WB	49.9	15.7	201.7	24.7	90.0	30.4
CTS Calgary	1st St. SW WB	34.6	11.1	176.6	19.6	80.0	39.9
CTS Calgary	3rd St. SW EB	40.0	16.2	181.4	22.1	80.0	28.9
CTS Calgary	City Hall EB	36.8	20.6	191.4	19.2	80.0	33.4
Muni San Francisco	Montgomery WB	34.4	11.0	146.0	23.6	60.0	29.6
NYCTA New York	Queens Plaza WB	40.7	17.3	134.7	30.2	53.0	6.4
NYCTA New York	Grand Central SB	64.3	16.7	164.7	39.0	53.0	14.1
NYCTA New York	Grand Central NB	53.9	14.8	184.1	29.3	53.0	47.5
PATH Newark	Exchange Place EB	23.3	7.4	115.8	20.1	55.0	22.6
PATH Newark	Journal Square WB	47.3	23.4	199.7	23.7	55.0	50.6
SkyTrain Vancouver	Broadway EB	30.2	2.6	145.6	20.7	40.0	70.2
SkyTrain Vancouver	Burrard WB	26.7	2.5	150.7	17.7	40.0	79.0
SkyTrain Vancouver	Metrotown EB	37.8	10.4	241.3	15.7	40.0	142.8
TTC Toronto	Bloor NB	43.0	15.3	145.5	29.4	55.0	17.0
TTC Toronto	King SB	28.1	5.9	168.3	16.7	55.0	73.4

Source: *TCRP Report 13* (*1*).
Note: NB: northbound, SB: southbound, WB: westbound, EB: eastbound.

Exhibit 8-17 shows the headway components graphically, with the operating margin as the end component of each bar. The bars are arranged in order of increasing headway. Note that the bar at the top is the only off-peak data set. It is included only for comparison and shows the large operating margin available when a system is not at capacity. The operating margins range widely and bear little relationship to system, technology, or loading levels.

Headway coefficient of variation.

A proxy for service reliability is the *headway coefficient of variation*—the standard deviation divided by the mean. There may be the expectation of a relationship between operating margin and service reliability; however, *TCRP Report 13: Rail Transit Capacity* (*1*) found no such relationship. Some inference can be drawn in that the system with the best headway adherence identified in *TCRP Report 13,* Vancouver's SkyTrain, also has the most generous operating margins.

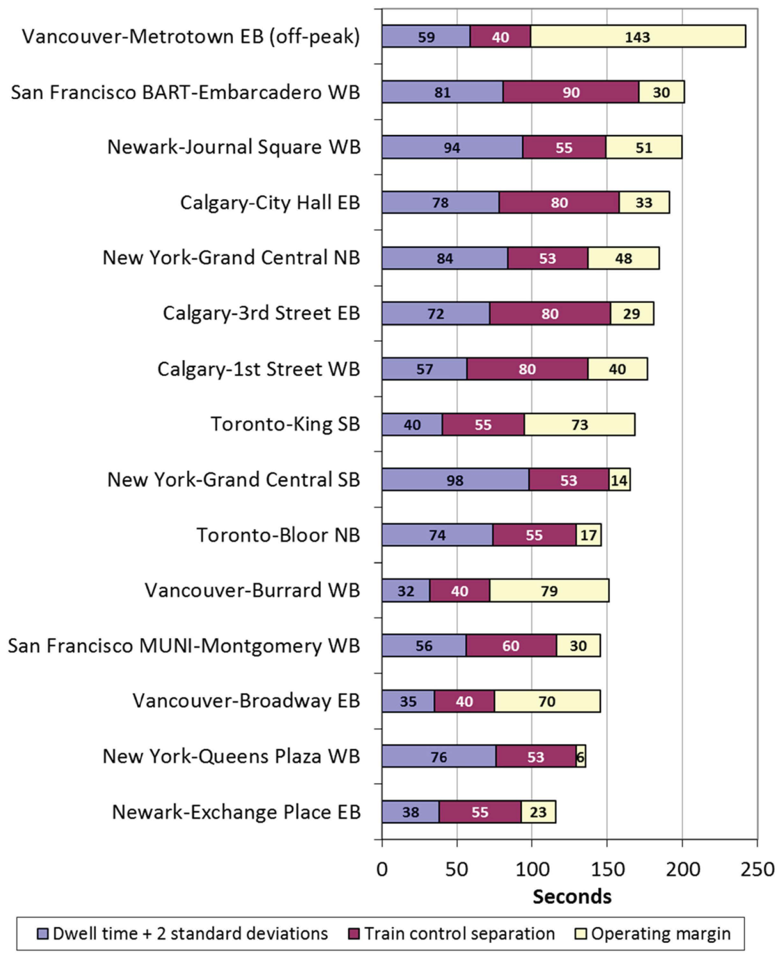


Exhibit 8-17 Headway Components of Surveyed Heavy Rail Transit Lines Operating at or Close to Capacity (1995)

Source: *TCRP Report 13* (*1*).
Note: NB: northbound, SB: southbound, WB: westbound, EB: eastbound.

Estimating Operating Margins

Although there is no clear relationship between existing rail transit operating margins and other operating criteria, this important factor, and the related terminal recovery or layover time, cannot be discounted. The inevitable headway irregularities and the need for reasonable operating flexibility require the greatest possible operating margin and recovery time to ensure reasonably even service and to achieve maximum capacity. Selecting a recommended operating margin is a dilemma, as too much reduces design capacity, but too little incurs sufficient irregularity that it may also serve to reduce capacity.

It is recommended that a range be considered for an operating margin. A reasonable level for a system with more relaxed loading levels, where all of the capacity is not needed, should be 35 s. On systems where headways prohibit such margin, a minimum level of 10 s can be used with the expectation that headway interference is likely.

Suggested operating margin range.

In between these extremes is a tighter range of 15, 20, or 25 s that is recommended. This range is used in estimating design capacity in this manual and is recommended as a default value for computations using the detailed procedures.

SKIP-STOP AND EXPRESS OPERATION

Skip-stop operation increases average speed but not capacity.

Skip-stop service (where a given train stops at every other station) is used on several of the heavy rail transit operations in Japan, New York, and Philadelphia. Skip stops provide faster travel times for most passengers, and require less equipment and fewer staff. They do not increase capacity as the constraint remains the dwell time at the maximum load point station at which all trains must stop. In fact, capacity can be slightly reduced as the extra passengers transferring between *A* and *B* trains at common stations can increase dwell times. Skip-stop operation is only applicable if the headways are sufficiently short that the "up to two-headway wait" at minor stations is acceptable to passengers.

Skip-stop service tries to balance loads between the "A" and "B" trains.

The common stations on the Japanese skip-stop operations have multiple platforms, typically two island platforms allowing passengers to transfer across the platform between *A* and *B* or between local and express trains.

Some light rail lines operate a form of express service, where a following train skips stops to catch up to—but not pass—the train in front.

Light rail operations may also skip stations when an *on-demand* operating policy is adopted. This requires that an onboard passenger signal to stop the train. Drivers must observe whether there are any waiting passengers as they approach each station. This is a particularly efficient way to increase line schedule speed and reduce operating costs. However, at higher capacity levels all trains will stop at all stations and so the practice has no effect on line capacity. Demand stops are rare on new North American light rail systems, even where there are clearly some low-volume stations where during off-peak times on-demand stops could contribute to lower energy consumption, lower maintenance costs, and a faster, more attractive service.

New York's Lexington Avenue express tracks carry 80% the number of trains of the local tracks, making it one of the most efficient four-track lines.

Directional capacity with express service.

Most trunk routes in New York have three or four tracks, while the Broad Street subway in Philadelphia and the North Side elevated in Chicago have four tracks. The capacity of four-track lines is not a simple multiple of two single tracks and varies widely with operating practices such as the merging and diverging of local and express services and trains holding at stations for local-express transfers. The result is that four tracks rarely increase capacity by more than 50% over a double-track line—and often less. A third express track does not necessarily increase capacity at all when restricted to the same station close-in limitations at stations with two platform faces.

PASSENGER-ACTUATED DOORS

Most new light rail systems have passenger-actuated doors, which increase comfort by retaining interior heat or air conditioning and reducing wear and tear on door mechanisms. The practice can extend station dwell time but is of little value at higher frequencies or busy stations where the use of all doors is generally required. Consequently, some systems use the feature selectively and allow the train operator to override passenger actuation and control all doors when appropriate.

Door cycle times.

A typical heavy rail transit car door will open and close in 5 s. Certain light rail doors associated with folding or sliding steps can take double this time to operate. A door opening initiated at the end of a station dwell will extend the dwell time by the door opening and closing time, plus any added passenger movement time. A system approaching its capacity could not tolerate such dwell extensions but would, in any event, be using all doors and might just as well be under driver control—avoiding any last-minute door opening and closing.

TRAIN AND PLATFORM SCREEN DOORS

Many AGT systems and newer heavy rail transit systems have platforms that are enclosed and separated by screens from the guideway, with doors that open directly adjacent to the train doors when the train stops at the platform. The door opening mechanisms are coordinated so that the platform and train doors open and close simultaneously. Examples include the Jubilee Line extension in London, underground stations on the Copenhagen Metro, the JFK AirTrain in New York, and virtually all on-airport people mover systems in the U.S. Platform doors provide superior safety and (potentially) climate control on the station platform by physically separating the platform zone from the trackway. Rapid transit systems in Hong Kong, Taipei, and Singapore use platform screen doors (some half-height, some full-height) and sometimes platform conductors at busier stations to control dwell time. Some platform screen doors are manually operated: the driver closes the platform doors first to shut off the crowds, and closes the train doors simultaneously if there is no obstruction, or holds the train doors open a little longer if needed to ensure everyone is safely inside the train. The platform conductor uses lightsticks and a whistle to indicate to passengers that train is ready to depart and the boarding process should stop.

Generally, these systems require automatic train operation to ensure that the train stops at exactly the right location on the platform for the train doors to line up with the platform doors. They also require uniform rolling stock with precisely-specified door locations.

FARE PAYMENT

Onboard light rail fare payment.

Fare payment can be a constraining factor on light rail systems that use manual on-board payment, although the trend for new systems is towards proof-of-payment fare collection systems that avoid the need for onboard fare payment. Manual fare payment adds an average of 1 to 3 s per boarding passenger. This is an inefficiency that increases running time, station by station, day by day.

However, the far more drastic impact of manual onboard fare collection is the restriction of boarding to a single staffed door. Not only do all passengers take more time to board individually, the efficiency of loading several passengers at once through multiple doors is lost, resulting in dwell times that are potentially three to four times as long as they would be without onboard fare collection. Exhibit 8-18(c) shows an extreme case of platform congestion resulting from manual onboard fare collection and a surge of passengers from a nearby tourist attraction. A system using manual onboard fare collection, and restricting boardings to driver-attended doors only, cannot achieve its maximum capacity.

A few systems provide onboard fare machines, combined with random fare checks. These machines allow passengers to board through all doors and then make their way inside the car to the fare machine. This addresses the dwell time issue, but can substitute a crowding and circulation issue inside the car, in the vicinity of the fare machine. On low-volume lines, or mainly pre-paid lines, congestion within the car may not be an issue. However, on high-volume systems with onboard fare machines, it can be an issue.

Exhibit 8-18
Onboard Fare Collection

(a) Farebox (Pittsburgh)

(b) Self-service machine (Portland Streetcar)

(c) Farebox, with passenger surge (Cleveland)

(d) Smart card reader (San Francisco)

Platform width.

Stations with high mixed flows must also have platforms of adequate width to accommodate the flows. Platform width is also a factor in making it easy for passengers to distribute themselves along the length of a train and to improve the peak-hour factor.

STATION AND PLATFORM DESIGN

Inadequate platform exit capacity can reduce line capacity, as dwell times of following trains increase.

Many station-related factors can influence demand. Poor location, inconvenient transfers to connecting modes, and inadequate or poorly located kiss-and-ride or park-and-ride facilities may all deter usage. However, the only factor that has a potential effect on the line capacity of a rail transit line is the rate of exiting from a platform. Adequate passageways, stairways, and escalators must be provided to ensure that a platform can clear before the arrival of the next train. Inadequacies in passenger access to a station may reduce demand but not capacity.

Chapter 10 of the TCQSM covers NFPA station exit requirements in greater detail.

Station exiting requirements are specified by the U.S. National Fire Protection Association (NFPA) Standard 130 for fixed-guideway transit systems (*6*). Exits, emergency exits, and places of refuge must be adequate to allow a platform with one headway's worth of passengers plus the entire complement of a full-length fully loaded train to be able to be evacuated to a safe location within 4 min. These regulations ensure that in all but the most unusual circumstances—where there is a disproportionate reliance on emergency exits—full capacity loads can leave the platform before the next train arrives.

NFPA 130 requirements may not be met on older systems. Additional exits must be provided to ensure that exit throughput is not constrained by platform back-ups. Rates of flow are established for passageways and for up and down stairs and escalators according to width. In emergencies, exit fare payment devices can be placed in a free passage mode. This is not the case in normal operation when adequate exit-fare control checks must be provided on those systems with distance-related fares.

Out-of-service fare gates and escalators have the potential to create station platform congestion if the remaining in-service equipment does not have sufficient capacity.

Even when NFPA 130 requirements are met, constraints posed by out-of-service fare gates, escalators, or other station components can potentially create congestion that could cause passenger queues to back up onto station platforms if these components are being operated close to their capacity. Chapter 10 of the TCQSM provides procedures for calculating the passenger-handling capacity of various station components.

WHEELCHAIR ACCOMMODATIONS

Since dwell times are one of the most critical components of headway, the time for wheelchair movements is important. Measured lift times run 2 to 3 min, with some as low as 60 s. The movement of wheelchairs on level surfaces is generally faster than walking passengers except where the car or platform is crowded. Level loading is essential to achieve high capacity. Where high platforms or low-floor cars cannot be provided, mini-high or high-block loading arrangements for wheelchairs, described later in this section, have the least impact on capacity. The vertical and horizontal gap between the edge of platform and the door is often a major problem for passengers in wheeled mobility aids.

An unknown is the number of customers in wheelchairs who will elect to use mainstream rail transit when all ADA measures have been implemented. A 1995 survey of heavily used rail transit systems indicated an average of 1 wheelchair use per 20,000 passengers (*1*). Other estimates range from 1 in 5,000 to 1 in 10,000. However, usage is usually dependent on other streetscape amenities and demographic factors as well. The usage of lifts can be three to five times higher than these rates due to use by other passengers not using wheeled mobility aids.

In addition to any boarding and alighting delays, the time for a wheelchair to move to a securement position and use any required securement or restraint systems can be considerable, particularly if the rail car is crowded. However, experienced users can be remarkably quick in boarding and alighting, and passenger movement times are often lower than for lift-equipped buses, as there is more room to maneuver wheelchairs, walkers, and scooters within rail vehicles. Off-vehicle fare collection also helps to speed loading for mobility limited and able-bodied passengers alike. The least loss of time occurs when the wheelchair position is close to the doorway and requires neither a folding seat nor the use of a securement system. Some systems have experienced passenger conflicts over mobility device seating priority when other passengers occupy the folding seats provided to create space for wheelchairs and other mobility devices.

Some agencies are overly cautious in adapting bus securement procedures to light rail service. Consideration of wheelchair securement is necessary for light rail vehicles operating on-street, due to the possibility of rapid braking as a result of traffic. However, many systems' experience indicates that wheelchair securement systems are not necessary for off-street rail service, as braking and acceleration is closely controlled and ride quality is smooth.

Impact of wheelchairs on line capacity.

There are many other types of boarding and alighting delays from passengers, other than those in wheelchairs, and generally these are accommodated in the operating margins and schedule recovery times. There is insufficient information to quantify the impact of wheelchair accessibility on line capacity. Indications are that in the short term, wheelchair lift and bridging plate use on light rail may cause delays, but this use is generally on systems with long headways (6 min and above) and have minimal impact on capacity at these levels. In the longer term, other accessibility requirements of the ADA and the move to level boarding with low-floor cars, or mini-high and profiled platforms, should sufficiently improve boarding and alighting movements to offset any negative impact of wheelchair use.

Wheelchair Boarding Methods

High-level loading is invariably used on heavy rail systems and automated guideway transit systems. The relative rarity of level loading with high-level platforms on other rail modes has resulted in a variety of methods to allow wheelchair access to rail vehicles. Each of the methods is outlined by mode in the sections that follow.

It should be noted that both mobility-impaired passengers and transit agencies prefer access methods that do not single out people with mobility impairments for special treatment. Lifts and special ramps cause delays that reduce the reliability of the service while isolating people with mobility impairments from other passengers. Mechanical devices such as lifts can also fail and put a train out of service. For these reasons, the popularity of lifts and other special devices for use by people with mobility impairments is decreasing in favor of more reliable and less exclusionary methods such as low-floor cars.

Light Rail

High Platforms

High platforms allow level movement between the platform and the car floor. This allows universal access to all cars of a train and removes the reliability and exclusionary effects associated with lifts, ramps, and special platforms. Passenger flow is sped up for all passengers since there are no steps to negotiate on the car. High-platform stations can be difficult to fit into available space, because of the need for an ADA-accessible sloped ramp to get between street level and platform level, which can increase costs. The use of high platforms on the transit mall portion of Calgary's light rail lines illustrates the difficulty accommodating this preferred loading method in on-street locations.

Mixed use of high and low platforms.

High platforms are also used at LRT stations in Buffalo, Pittsburgh, and San Francisco in combination with low-level loading at other stops. Buffalo is unusual in that a subway with high platforms serves the outer portion of the line, while the downtown segment is on a transit mall with low-level loading using fold-out steps combined with high-platform stubs for wheelchair access. Pittsburgh has separate doors for each platform level, while the San Francisco Muni uses cars fitted with steps that can be mechanically raised to floor height at high-platform stations.

Examples of high-platform stations and vehicles used in mixed high- and low-platform environments are shown in Exhibit 8-19.

Exhibit 8-19
High-Platform Station and Adjustable Door Height Examples

(a) On-street station in median (San Francisco)

(b) Partial platform (Buffalo)

(c) Adjustable vehicle steps (San Francisco)

(d) Separate door levels (Pittsburgh)

Low-Floor Cars

Low-floor cars improve access for everyone, not just persons using mobility aids.

Low-floor cars offer a straightforward solution to the need for universal access to light rail vehicles. By bringing the floor height down to just above the railhead, boarding is simplified for all passengers, as steps are no longer required. Small, extendible ramps and slight increases in platform edge height allow passengers using wheeled mobility aids to board without the aid of lifts or special platforms. Boarding by persons with strollers, bicycles, and luggage, and by persons who have difficulty climbing steps is also greatly simplified. Exhibit 8-20 presents examples of low-floor light rail cars.

Exhibit 8-20
Low-Floor Car Examples

(a) Portland, Oregon (light rail)

(b) Jersey City, New Jersey

Low-floor cars provide much of the benefit of level loading without the need for high platforms. The typical floor height is 14 in. (350 mm), about double the height of a normal curb. Medium- or intermediate-height platforms are therefore still required for no-step boarding, but long ramps are unnecessary. Buttons located at a lower height than the separate passenger-actuated door buttons on the inside and outside of the car allow wheelchair users to deploy the ramp on demand to bridge the gap between the car and the platform.

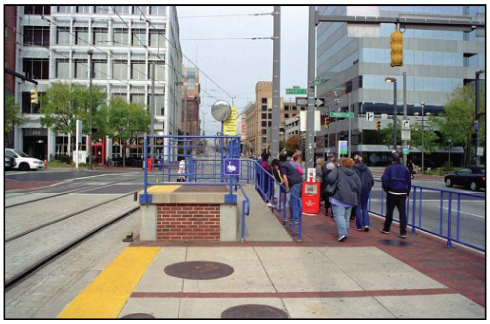
While low-floor cars have operated in Europe since the mid-1980s, the first North American operation began on Portland's light rail system in 1997. Portland's low-floor cars are compatible with the system's older high-floor fleet, allowing two-car trains to be formed from one high-floor and one low-floor car. Low-floor cars have subsequently been placed in service on many new North American light rail lines.

Mini-High Platforms

The most common wheelchair access methods to high-floor light rail cars are *mini-high* platforms that provide level loading to the accessible door of the train. This method is mechanically simple and often uses a folding bridgeplate, manually lowered by the train operator, to provide a path over the stepwell between the platform edge and vehicle floor. The mini-high platform is reached by a ramp or, where space limitations require, a small lift. A canopy is sometimes provided over the ramp. Exhibit 8-21 provides examples of mini-high platforms used in North America.

Exhibit 8-21
Mini-High Platforms

(a) San Francisco

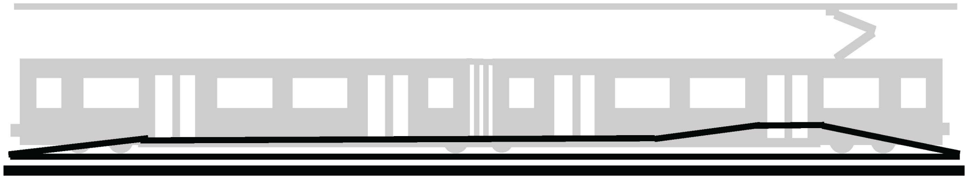


(b) Baltimore

Profiled platform as an alternative to mini-high platforms.

An alternative to the mini-high platform is the Manchester-style profiled platform, shown in Exhibit 8-22. This platform has an intermediate height and is profiled up to a section that is level with one doorway for wheelchair access. Maximum platform slopes are shown in Exhibit 8-23.

Exhibit 8-22
Profiled Light Rail Platform Providing for One Accessible Door



Source: *TCRP Report 13* (*1*).

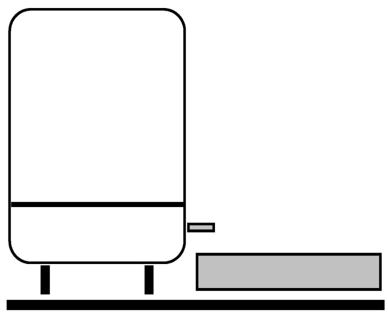
Exhibit 8-23
ADA Maximum Platform Slopes

Maximum Rise		Maximum Slope	
Rise ≤ 3 in.	Rise ≤ 7.5 cm	1:4	14.04°
3 in. < Rise ≤ 6 in.	7.5 cm < Rise ≤ 15 cm	1:6	9.46°
6 in. < Rise ≤ 9 in.	15 cm < Rise ≤ 22.5 cm	1:8	7.13°
Rise > 9 in.	Rise > 22.5 cm	1:12	4.76°

Folding steps and profiled platforms.

Most of the platform is only slightly higher than a sidewalk. Where the street arrangement permits, the profiled platform can be raised so that its mid-section—taking up most of the length—is raised one step, providing single-step entry to most doors. Alternatively, cars can have a slide-out or fold-down step as shown in Exhibit 8-24.

Exhibit 8-24
Profiled Light Rail Platform with Slide-Out or Fold-Down Step



Source: *TCRP Report 13* (*1*).

Car-Mounted Lifts

Car-mounted lifts, illustrated in Exhibit 8-25, were introduced on the San Diego Trolley, one of the first light rail transit systems to be wheelchair accessible. In San Diego, lifts are mounted in older high-floor cars so that the first door on the right side of every train is lift-equipped. When not in use, the lift is stored in a vertical position that blocks the doorway from use by other passengers. The Kenosha, Wisconsin, vintage trolleys use a car-mounted lift that folds flat against the side of the door when not in use, which allows other passengers to use the door when the lift is not in use. Trains used on New Orleans' Waterfront and Canal Street lines use car-mounted lifts located at a high-level door not used by other passengers.

Dwell times with car mounted lifts.

Boarding and alighting times with the car-mounted lifts are around 1 min for each passenger movement. However, the need for the train operator to leave the cab to operate the lift adds to the time required and can mean the total station dwell time extends to 1.5 to 2 min when the lift is used. If the operator is required to assist in securing the wheelchair, the dwell can be further extended.

Exhibit 8-25
Car-Mounted Lifts

(a) San Diego (b) Kenosha (c) New Orleans

Platform-Mounted Lifts

Platform mounted lifts were originally used on the San Jose and Portland light rail systems. They offer advantages over car-mounted lifts in that all car doors are left available for other passengers when the lift is not required, the lift is not subject to car vibration, and the failure of a lift need not remove a car from service. Disadvantages include increased susceptibility to vandalism and an increase in the distance that the train operator must walk to operate the lift. Wheelchair boardings using platform lifts took 2 to 3 min, giving a total train delay (including loading and unloading) of 4 to 6 min per passenger requiring the lift. These delays could easily consume a train's scheduled terminal recovery time. Both systems removed their platform lifts after introducing low-floor cars to their fleet.

Commuter Rail

Bridgeplates are often used to span the gap between platform and train.

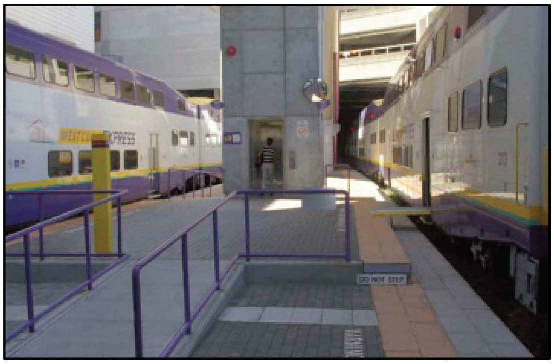
Commuter rail systems use many of the same kinds of access methods as light rail systems. The main difference is that these methods are often supplemented with a bridgeplate to span the gap between platform and train when a form of level boarding is used. The vertical and/or horizontal gap between the train and platform for "level" boarding typically is greater for commuter rail than for light rail. The bridgeplate can be portable or built into the train.

High-level platforms are usually not possible on lines shared with freight trains.

High-level platforms provide the easiest and fastest boarding for all passengers. The Electric Division of Chicago's Metra, MTA-Long Island Rail Road, and MTA-Metro North Railroad are among the commuter rail lines that use high-level platforms. However, it is often not possible to provide high-level platforms on lines that are shared with freight trains, as freight wide-loads will need to be accommodated (*1*).

Mini-high platforms, combined with a bridgeplate (Exhibit 8-26), are a frequently used option on lines with low-level platforms. Platform lifts are used by CalTrain in the San Francisco Bay Area and at some New Jersey Transit commuter rail stations. Metra's diesel-powered lines in Chicago provide cars with onboard lifts.

Exhibit 8-26
Commuter Rail Wheelchair Loading Examples



(a) Vancouver

(b) Seattle

Ropeway Modes

Inclined Planes

Because each inclined plane in North America is unique, so are the means of providing access; due to their age, many inclined planes are not accessible as a result of the vehicle and/or station design.

Access is much easier to provide when the car floor is level, rather than when the seats are tiered (as is the case on most inclined planes). Johnstown, Pennsylvania, has level loading from each end. The Horseshoe Curve funicular, near Altoona, Pennsylvania, provides level loading from each side of the car.

Several access methods have been developed for tiered cars. The funicular at the Industry Hills Resort in California was designed to carry golf carts and has a series of terraced ramps leading to each car tier. Pittsburgh's Monongahela Incline has an elevator inside the lower station to take wheelchairs to the top tier loading area; wheelchairs exit on the level at the upper station. Los Angeles' Angels Flight uses an inclined platform lift (like those used on stairways) to bring wheelchairs to the top car level at the lower station; wheelchairs exit on the level at the upper station.

Aerial Ropeways

In the past, gondola access required that the entire system be brought to a stop to load and unload wheelchairs because boarding normally occurred as the carriers circulated through the station while moving (typically at 50 ft/min or 15 m/min) and there was a vertical gap between the cabin floor and the platform that needed to be overcome. Newer designs provide a trench in the platform floor that the gondola passes through, allowing level loading. Clutching equipment allows an individual carrier to be brought to a near-stop to load wheelchairs, without stopping the entire system. Aerial tramways provide level boarding from the platform into the cabin; however, elevators or ramps may be needed to access the platform.

SYSTEM DESIGN

Although the capacity estimation procedures in Section 5 are focused on normal operating conditions, it is prudent to consider the impacts of abnormal conditions. Three areas in particular to consider are (a) the potential impacts of disabled trains on system operation, (b) routine track maintenance, and (c) the handling of special event crowds.

Disabled Trains

When a train needs to be taken out of service, it is desirable to get it off the main line as quickly as possible. In a typical two-track operation, this means moving it to a place where it can be temporarily stored, moving it into a yard, or turning it back on the opposite track. The train will likely not be able to move at its normal speed, which means that the trains following it will catch up to it and be delayed. The longer the disabled train remains on the line, the more trains will be delayed. Because the disabled train, and the ones delayed behind it, will occupy signal blocks for longer periods of time, train headways will increase and line capacity will decrease. The longer headways mean that more passengers will accumulate on platforms between trains, potentially leading to platform crowding. In addition, passengers on a disabled train will need to be off-loaded; the station where they off-load will have passengers of its own on the platform waiting to board the train, and the platform should be designed to accommodate all of those passengers. Finally, when a system operates close to capacity, any significant delay will use up a train's terminal recovery time, resulting in potential delays later on in the reverse direction. As can be imagined, a disabled train can quickly cause delays and crowding that ripple along the line, and which may take a long time to clear up.

Crossovers can be used to short-turn trains to help recover from delays.

Exhibit 8-27 shows examples of ways that a rail system can be designed to accommodate disabled trains. Trains can be stored off the main line in a pocket track accessible to both main tracks (allowing the train to be reversed later if needed), on a siding accessible to only one direction, or in a storage track beyond a terminal station. If a yard is convenient to the train's location, it can be removed from the line altogether. Finally, a train can be turned back at a crossover—either to get to a storage track or to move it in a direction where it will delay fewer people. Crossovers can also be used to short-turn trains in advance of their scheduled terminus—this can help to fill a gap in the sequence of trains in the reverse direction, helping to reduce the time needed to recover from the delay, at the expense of further delaying passengers traveling beyond the station at which the train makes its short turn.

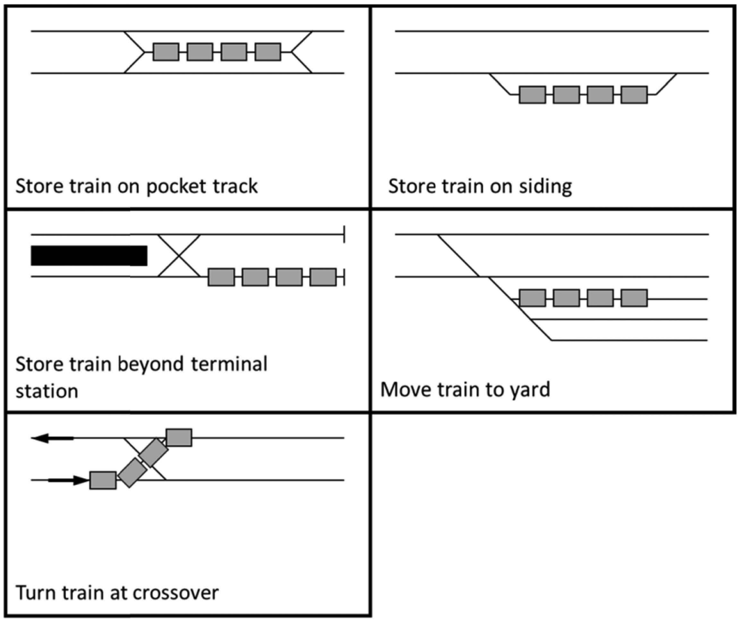


Exhibit 8-27 System Design Features for Accommodating Disabled Trains

The spacing of storage tracks and crossovers requires balancing initial capital costs when constructing the system with the amount of delay a system is willing to tolerate when a train breaks down. Physical constraints, particularly when tracks are elevated or underground, must also be considered. Train headways and the resulting line capacity as a result of a disabled train can be estimated from agency experience, or by using an assumed disabled train operating speed, the resulting block traversal time, and increased dwell times for subsequent trains resulting from the longer headways and greater passenger accumulations in stations.

Guidance on estimating line capacity with a disabled train.

Track Maintenance

Many rail systems do not operate 24 hours a day because ridership levels do not warrant it. Overnight closure provides a window of time to conduct routine track maintenance when the tracks are out of service (e.g., from 1 a.m. to 4 a.m.). However, some projects may require more time than this window allows, or an agency may have a need for 24-hour operation. An alternative means of moving passengers must be developed when a track needs to be taken out of service during regular service hours.

If passenger demand is low, the remaining in-service track can be used in single-track operation to move trains around the work area if signaling is provided for the wrong-side direction. However, the capacity of both directions will be greatly reduced. The single-track capacity procedures presented in Section 5 can be used in these circumstances, given a known distance between crossovers.

Single-track rail operation.

If single-track operation does not provide sufficient capacity to meet passenger demand, another alternative is to provide a bus bridge. In this situation, trains are turned back on either side of the work area, and passengers transfer to buses to meet a train on the other side of the work area or to reach a destination station within the work area.

Bus bridges.

Use of express tracks.

MTA-New York City Transit is able to take advantage of the third and fourth tracks that exist on many of its major lines to close tracks for maintenance and maintain two-direction operations.

Special Events

Special events—such as sporting events, concerts, and community festivals—can generate very large passenger demands during a short span of time. While passengers are willing to tolerate longer delays and greater levels of crowding under these circumstances than they might otherwise, system design should still consider any special train storage needs in order to make sure that crowds can be transported away from the event site in a reasonable period of time. San Diego, for example, designed its football stadium station with storage for 18 cars—enough room for five three- to four-car trains. For Super Bowl XXXII, San Diego closed the Mission San Diego station, at the time a terminal station located one station east of the stadium, which allowed storage on the main tracks for twenty-one three- to four-car trains. Light rail was able to transport 29,800 passengers—30% of the Super Bowl's attendance—within 2 hours following the end of the game (*7*).

Crowd management.

Crowd management is another issue requiring consideration. Security personnel are usually needed to keep passengers off tracks and to limit platform access to avoid overcrowding problems. Providing pre-sold return tickets and/or providing mobile ticket sales outlets minimizes crowds and delays at ticket machines. Platforms should be sized to accommodate expected special event crowds, and additional temporary space may be required to queue passengers when there are constraints on platform space. For example, Muni's 2nd & King light rail station, adjacent to San Francisco's baseball stadium, is located in a street median and has little platform room for large event crowds. Instead, passengers are queued using portable fences in the adjacent closed-off street following games and are allowed onto the platform when a train arrives. San Diego's football stadium station has three platforms, allowing trains to be loaded from both sides, minimizing dwell time.

Demand management.

Demand management measures can be used to spread out passenger demand following sporting events, and thus minimize platform crowding. During the sold-out first season at their new baseball stadium, for example, the Seattle Mariners provided post-game trivia contests and a ceremonial closing of the stadium's retractable roof to encourage a portion of the fans to linger after the game.

5. RAIL SYSTEM CAPACITY METHODOLOGIES

INTRODUCTION

Except for the simplest of operations, or for operations that have a long history of experience, there are no "one-size-fits-all" formulas for calculating the capacity of a rail corridor. There are many variables involved (train characteristics, speed limits, train and siding lengths, signal system characteristics, etc.), which means that a simple approach may not yield precise and accurate results. Rules of thumb and deterministic formulas that estimate minimum headways and capacities based on operational performance measures provide useful surrogates that can be used to estimate the capacity of a line or system for planning purposes. However, complex networks, terminals, and systems with complicated train operating patterns are difficult to analyze using sketch methods because of their complexity and the fact that increasing operational and physical complexity tends to either reduce the capacity of each individual element of the network or reduce the reliability of the system when it is operated close to its capacity. As a result, capacity analyses for complex networks and operating plans, where a high level of accuracy is required, are able to utilize detailed operations simulation models and methods. These are discussed in Section 6.

Methodologies presented in this section are suitable for planning purposes. Simulation is typically used for more detailed operations analyses.

This section is divided into four main subsections: a general methodology suitable to many types of rail transit operations (e.g., heavy rail, light rail, and commuter rail not sharing trackage with non-commuter trains); a discussion of the challenges of determining commuter rail capacity using deterministic methods when tracks are shared with freight or intercity passenger trains; guidance on adapting the general methodology for AGT; and a methodology for determining the capacity of ropeway modes.

GENERAL METHODOLOGY

Overview

The general methodology seeks to identify the weakest link along the rail line that will ultimately control train throughput by setting the minimum headway that can be operated between trains. The following are the potential weak links:

- Dwell time at the controlling station along the line;
- Minimum train separation allowed by the train control system;
- Right-of-way characteristics (e.g., single-track operation);
- Turnbacks;
- Junctions;
- Power supply constraints;
- Train length limitations; and
- Track configuration within terminals.

On a new system, the sum of the controlling dwell time, the minimum train separation, and the operating margin will typically control the line capacity. However, with older systems or with newer systems built to minimize initial costs, one of the other factors may turn out to control capacity. Newer systems can often be designed to accommodate future projects to eliminate these initial constraints, but it may be costly to upgrade older systems for which allowances were not originally made to address built-in capacity constraints.

A spreadsheet that implements the general rail capacity methodology is provided on the accompanying CD-ROM.

Consider using the planning graphs in Section 6 when input variables must be defaulted.

The general methodology can be applied using hand calculations or by using the spreadsheet on the accompanying CD-ROM. When there is uncertainty about input values to these equations or where several of the performance variables are unknown (e.g., the technology or specific vehicle has not been selected), then the use of this procedure is not recommended. The planning graphs found in Section 6 provide *generic design capacity ranges* with less effort and potentially as much accuracy as the general methodology in which one or more input factors will have to be estimated.

Step 1: Determine the Non-Interference Headway

Step 1a: Determine the Maximum Load Point Station

Traditionally, the maximum load point station is the principal downtown station, or the downtown station where two or more rail transit lines meet. However, this is not always the case. With increasingly dispersed urban travel patterns, some rail transit lines do not serve the downtown. Los Angeles' Green Line and extensions to Vancouver's SkyTrain are examples.

Ridership models.

A regional transportation model will usually produce ridership data by station, both boardings and alightings and direction of travel. Such data are usually for a peak-hour or a multiple-hour peak period and rarely for the preferable 15-min period. Depending on the number of zones and nodes in the model, data accuracy at the station level can be poor—particularly if there is more than one station in a zone. Nevertheless, this is often the sole source of individual station volumes, and without it the selection of the maximum load point station requires an educated estimate. Present and future CBD employment sometimes can be used to estimate the reasonableness of ridership forecasts.

Step 1b: Determine the Control System's Minimum Train Separation

This step determines the minimum train separation associated with three types of train control systems, each providing progressively increased throughput:

- Three-aspect fixed-block signaling system,
- Multiple-command cab signaling, and
- Moving-block signaling system.

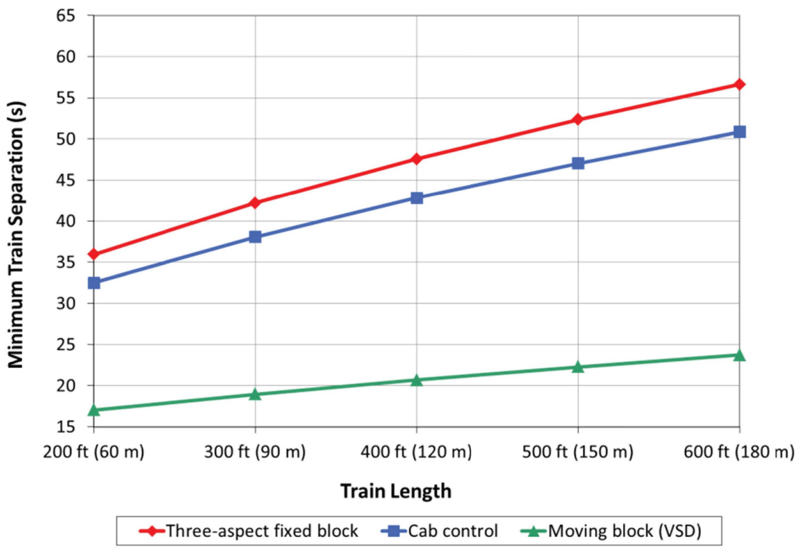
Although the equations that follow appear long, the arithmetic is simple and can be implemented using basic functions in a spreadsheet. However, before going to this effort, check the availability of the required input parameters in Exhibit 8-28. Parameters can be adjusted for system specific values or left at their default value. Train length is the most important variable. If most parameters are left at their default values, it would be easier to refer to Exhibit 8-29, which shows the minimum train control separation against train length for the three types of train control systems.

Exhibit 8-28
Minimum Train Control Separation Parameters

Default Value	Term	Description
Calculated	t_{cs}	train control separation (s)
650 ft, 200 m	L_t	longest train length (ft, m)
35 ft, 10 m	d_{eb}	distance from the front of stopped train to start of station exit block (ft, m)
Calculated	v_a	station approach speed (ft/s, m/s)
88 ft/s, 27.8 m/s	v_{max}	maximum line speed (88 ft/s = 60 mi/h, 27.8 m/s = 100 km/h)
75%	f_{br}	braking safety factor—worst-case service braking is f_{br}% of specified normal rate—typically 75% (decimal)
2.4–three-aspect, 1.2–cab, 1.0—moving block	b	separation safety factor—equivalent to number of braking distances (surrogate for blocks) that separate trains
3.0 s	t_{os}	time for overspeed governor to operate on automatic systems—to be replaced with driver sighting and reaction times on manual systems
0.5 s	t_{jl}	time lost to braking jerk limitation
1.5 s	t_{br}	brake system reaction time
4.3 ft/s^2, 1.3 m/s^2	a	initial service acceleration rate (ft/s^2, m/s^2)
4.3 ft/s^2, 1.3 m/s^2	d	service deceleration rate (ft/s^2, m/s^2)
32 ft/s^2, 10 m/s^2	a_g	acceleration due to gravity (ft/s^2, m/s^2)
0%	G_i	grade into station, downgrade = negative (decimal)
0%	G_o	grade out of station, downgrade = negative (decimal)
90%	l_v	line voltage as percentage of specification (decimal)
20.5 ft, 6.25 m	P_e	positioning error—moving block only (ft, m)
165 ft, 50 m	S_{mb}	moving-block safety distance—moving block only (ft, m)

Source: *TCRP Report 13* (*1*).

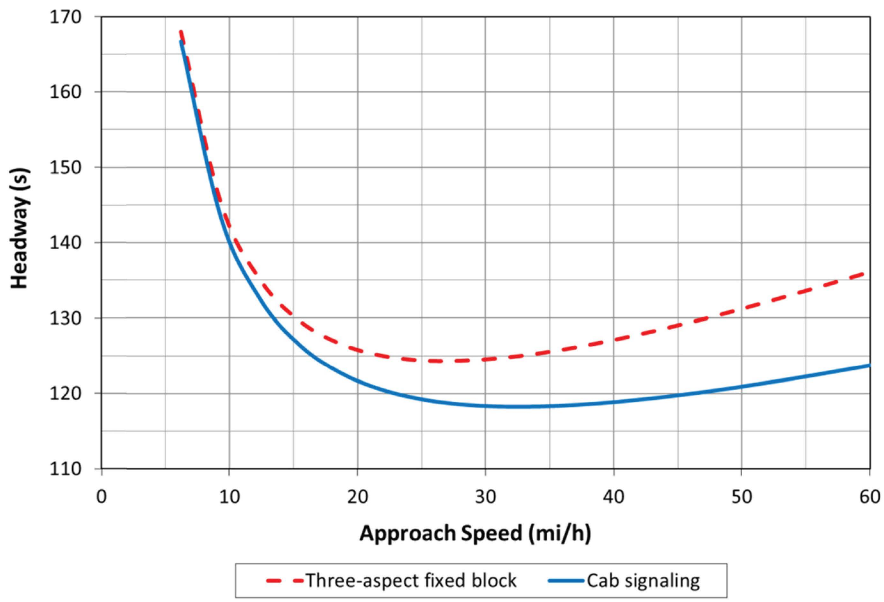
Exhibit 8-29
Typical Minimum Train Separation Versus Train Length



Source: *TCRP Report 13* (*1*).
Note: VSD = variable safety distance.

Exhibit 8-30 shows minimum station headways achieved using the typical values shown in Exhibit 8-28, derived from the equations presented later in this section, and including an assumed dwell time and operating margin. The optimum approach speeds shown in this exhibit should be compared with the maximum speeds imposed by switches and curves in the vicinity of the maximum load point station.

Exhibit 8-30
Typical Station Headways for Lines at Capacity

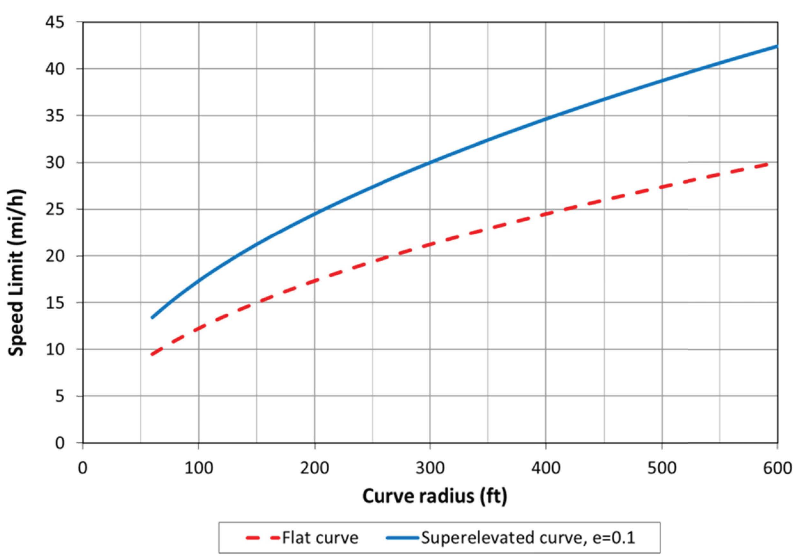


Source: *TCRP Report 13* (*1*).
Note: Assumes 45-s average dwell time and 20-s operating margin.

Exhibit 8-30 shows that the optimum approach speed for three-aspect fixed-block signaling in this situation is 28 mi/h (45 km/h), while, for cab signaling, the optimum approach speed is 32 mi/h (52 km/h). If special work (interlockings) or curves restrict approach speeds below these values, then the lower values must be calculated and used. Typical speed limits for curves and turnouts (switches) are shown in Exhibit 8-31 and Exhibit 8-32, respectively. Determine any such station approach speed restrictions and their distance from the station stopping point. Next, compare this speed restriction with the normal approach speed at that distance from the station as shown in Exhibit 8-33. The most restrictive approach speed is used in the equations presented in this section.

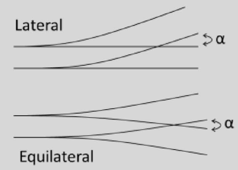
Exhibit 8-31
Maximum Speed Limits on Curves

Curves and turnouts (switches) impose speed restrictions.



Source: *TCRP Report 13* (*1*).
Notes: Transition spirals are not taken into account. A metric version of this exhibit appears in Appendix A.

Exhibit 8-32
Speed Limits on Turnouts

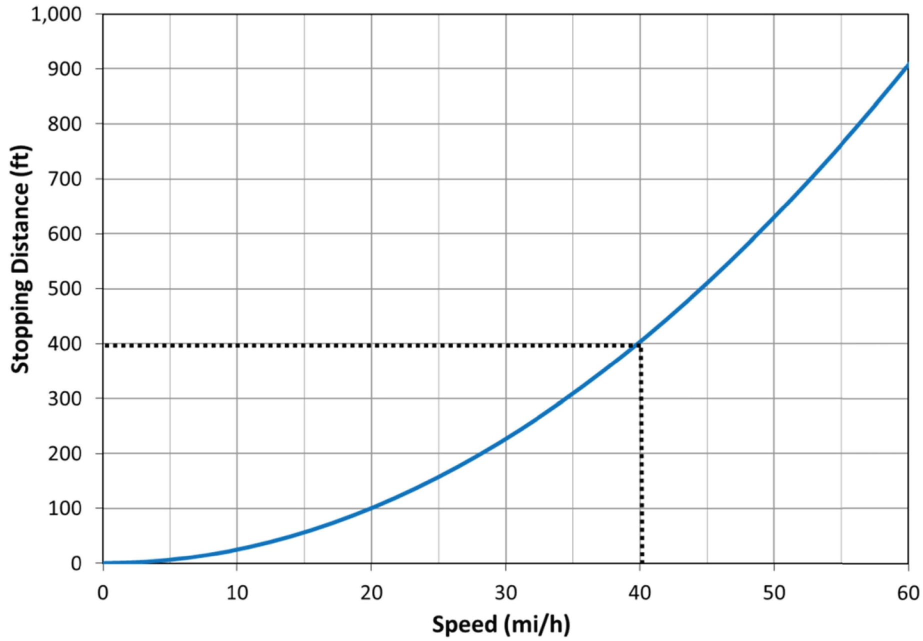


Turnout Number	Lateral Turnout mi/h	Lateral Turnout km/h	Equilateral Turnout mi/h	Equilateral Turnout km/h
#6	15	24	21	34
#8	20	32	28	45
#10	25	40	35	57
#20	50	81	70	113

Source: *AREMA Manual for Railway Engineering* (*8*).
Notes: Speeds shown are based on freight trains using level turnouts with curved switch points. Intercity passenger and rail transit cars are designed for greater roll through curves and can operate comfortably at somewhat higher speeds than shown. Many transit agencies have their own speed limits for turnouts that differ from those shown.

Exhibit 8-33
Typical Stopping Distance as a Function of Speed

Use this chart to find how far from the station a computed optimal approach speed will occur, then determine if there is a lower speed limit at that location.

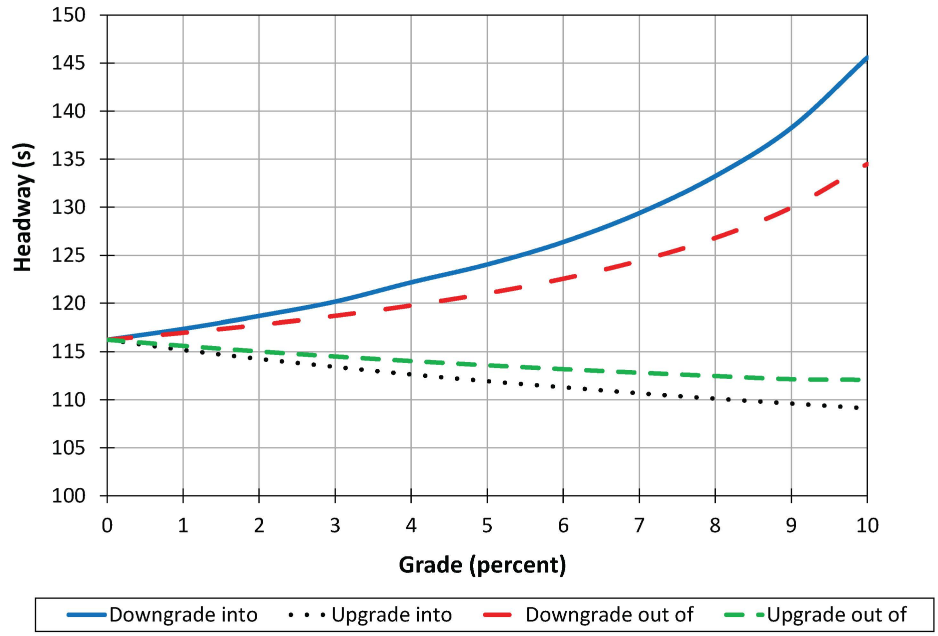


Source: *TCRP Report 13* (*1*).
Note: A metric version of this exhibit appears in Appendix A.

The dotted line example in Exhibit 8-33 shows that at 400 ft (120 m) from a station, expressed as the distance from the front of the approaching train to the stopping point, the approaching train will have a speed of 40 mi/h (64 km/h). If there is a speed limit at this point that is lower than 40 mi/h (64 km/h), then the minimum train separation, t_{cs}, must be calculated with the approach speed v_a set to that limit.

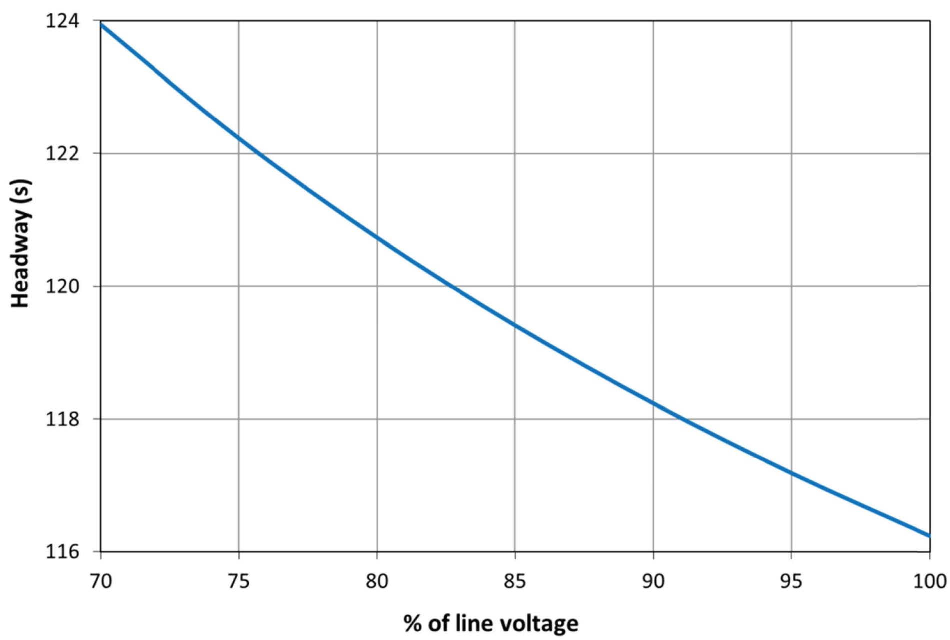
Two other factors affect minimum headways. Grades into or out of a station will change the acceleration and braking rates. Line voltage will drop below the nominal value on heavily used systems and reduce train performance. The results of grades and voltage drops are shown in Exhibit 8-34 and Exhibit 8-35, respectively. The calculations of these effects are complex and best left to a computer simulation. If a simulation model is not available, then the approximate headway changes can be read from Exhibit 8-34 and Exhibit 8-35, and the calculations adjusted by the appropriate number of seconds.

Exhibit 8-34
Typical Effect of Grade on Station Headway



Source: *TCRP Report 13* (*1*).
Note: Assumes cab signals, 45-s dwell time, and 20-s operating margin.

Exhibit 8-35
Typical Headway Changes with Voltage



Source: *TCRP Report 13* (*1*).

Fixed-Block and Cab Signaling Throughput

The minimum train control separation for fixed-block and cab signal systems is given by Equation 8-1, with variables as shown in Exhibit 8-28:

$$t_{cs} = \sqrt{\frac{2(L_t + d_{eb})}{a + a_g G_o}} + \frac{L_t}{v_a} + \left(\frac{1}{f_{br}} + b\right)\left(\frac{v_a}{2(d + a_g G_i)}\right) + \frac{(a + a_g G_o) l_v^2 t_{os}^2}{2v_a}\left(1 - \frac{v_a}{v_{max}}\right) + t_{os} + t_{jl} + t_{br}$$

Equation 8-1

This equation should be solved for the minimum value of t_{cs}. The approach speed v_a that produces this minimum value must then be checked against any speed restrictions approaching the station from Exhibit 8-33.

Compare the approach speed producing the minimum train separation to any speed restrictions on the station approach.

Moving-Block Throughput

Moving-block signaling systems replace separation by fixed blocks with a moving block based on the braking distance to a target point plus a safety separation distance. The safety separation distance can be fixed for a given system and type of rolling stock or can be continually adjusted with speed and grades.

Moving-block train separation safety distances can be fixed or variable.

Equation 8-2 determines the train control separation for a moving-block signaling system with fixed safety separation, with variables as given in Exhibit 8-28. Note that the time for the overspeed governor to operate is incorporated into the safety distance and so does not appear in the equation.

$$t_{cs} = \frac{L_t + S_{mb}}{v_a} + \frac{1}{f_{br}}\left(\frac{v_a}{2d}\right) + t_{jl} + t_{br}$$

Equation 8-2

Note that this equation is not affected by either line voltage or station grade. Lower voltages increase the time for a train to clear a station platform. In moving-block systems this time does not affect throughput. When a train starts to leave a station, the target point of the following train is immediately advanced accordingly. The worst-case approach grade is included in the determination of the safety distance. This can result in sub-optimal minimum train separation.

Higher throughput is usually obtained with a moving-block signaling system with a variable safety distance consisting of the braking distance at the particular speed plus a runaway propulsion allowance. The minimum train control headway of such a system is given by Equation 8-3, with variables and default values as given in Exhibit 8-28.

$$t_{cs} = \frac{L_t + P_e}{v_a} + \left(\frac{1}{f_{br}} + b\right)\left(\frac{v_a}{2(d + a_g G_i)}\right) + \frac{(a + a_g G_o) l_v^2 t_{os}^2}{2v_a}\left(1 - \frac{v_a}{v_{max}}\right) + t_{os} + t_{jl} + t_{br}$$

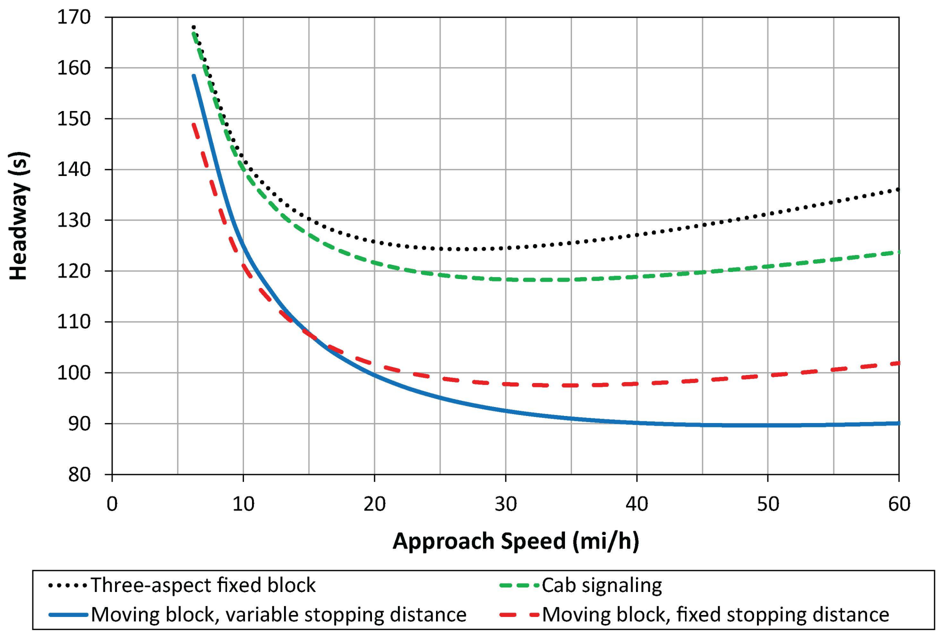
Equation 8-3

Equation 8-3 adjusts the safety separation entering a station due to any grade. A downgrade will increase the braking distance and so require a longer safety separation—and vice versa.

The results of Equation 8-2 and Equation 8-3 are shown in Exhibit 8-36. The resultant minimum station headway of 97 s occurs at an approach speed of 35 mi/h (56 km/h). The respective curves for a conventional three-aspect fixed-block signaling system and a cab signaling system are included for comparison. As would be expected, a moving-block system with a speed variable safety distance shows the lowest overall headway. The difference between the two methods of determining the safety distance represents an 8-s difference in the minimum headway. Voltage fluctuations have little effect on moving-block headways as the time to clear the platform is not a component in calculating the moving-block signaling system headway.

Exhibit 8-36
Typical Moving-Block Station Headways Compared with Conventional Fixed-Block Systems

Capacity is higher with a moving-block signaling system.



Source: *TCRP Report 13* (*1*).
Notes: Assumes 45-s dwell time and 20-s operating margin. A metric version of this exhibit appears in Appendix A.

The appropriate equation above should be solved for the minimum value of t_{cs}. The approach speed, v_a, that produces this minimum value must then be checked against any speed restrictions approaching the station from Exhibit 8-33.

Check Results

Compare the approach speed producing the minimum train separation to any speed restrictions on the station approach.

Compare the results obtained from the above equations with Exhibit 8-29. The calculated minimum train separation should be close to or moderately greater than the values charted. If lower, there is probably an error, as the charted values are the minimums using typical maximum rail transit performance criteria without applying any corrections for grades or speed restrictions into or out of the station.

Step 1c: Determine the Average Dwell Time at the Critical Station

The train close-in time at the critical station depends on a train's physical performance and length, as well as other fixed-system characteristics, and therefore can be calculated with some precision. Station dwell time cannot be determined with the same exactitude. Virtually all the literature references related to rail transit capacity assigns a set time to dwell time. Many simulation models do likewise, using typical figures of 15 to 20 s for lesser stations and 30 to 45 s for major stations. The main constituents of dwell time are as follows:

- Passenger flow time at the busiest door,
- Remaining (unused) door-open time, and
- Waiting to depart time (with doors closed).

Four methods of estimating controlling dwell.

Four methods of estimating dwell time or controlling dwell time are provided in this section. The first method is the one used in most of the literature references—simply assigning a reasonable figure to the critical station. The second method uses field data reported in *TCRP Report 13* (*1*), allowing the selection of a controlling dwell time from the critical station of rail transit lines with similarities to the one being analyzed. These two methods are suitable where information on passenger flows at the critical station is not available.

The third method is only suitable for new lines in cities with existing rail transit systems. In this case, using the mean dwell time plus two standard deviations is suggested, based on a comparable station on the existing network. The fourth and final method uses a statistical approach of determining station dwell times based on peak-hour passenger flows. This method, detailed in *TCRP Report 13* (*1*), is complex and still requires an estimate of the ratio of the busiest door to average door flow.

None of these methods are entirely satisfactory. This explains why practitioners over a period of several decades have resorted to simply assigning a reasonable value to station dwell time.

Method 1: Assigning a Value

Existing rail transit systems operating at or close to capacity have median station dwell times over the peak hour that range from 30 to 50 s with occasional exceptional situations—such as the heavy peak hour mixed flow at NYCT's Grand Central Station of more than 60 s. A tighter range of dwell time values—35 to 45 s—was used to develop the capacity graphs that appear in Section 6.

Method 2: Using Existing Dwell Time Data

Examples of existing dwell time data from the highest-use station on lines that are close to capacity are summarized in Exhibit 8-37. Selection of a dwell time from this table is less arbitrary than Method 1 and allows some selectivity of mode and the opportunity to pick systems and stations with similar characteristics to those of the one under examination.

The selected median station dwell times range from 27.5 to 61.5 s. The highest values are mainly alighting and mixed-flow records from manually operated systems with two-person crews. Most station dwell times in Exhibit 8-37 fit into the 35 to 45 s range suggested in the previous method.

Exhibit 8-37
Peak-Period Station Dwell Times for Heavily Used Systems (1995)

System & City	Station	Passengers Observed	Mean Dwell (s)	Mean Headway (s)
BART (San Francisco)	Embarcadero WB	2,298	48.0	155.0
CTS (Calgary)	1st St. WB (LRT)	298	33.0	143.0
CTS (Calgary)	3rd St. WB (LRT)	339	38.0	159.0
CTS (Calgary)	City Hall EB (LRT)	201	34.0	161.0
NYCT (New York)	Grand Central SB (4&5)	3,488	61.5	142.5
NYCT (New York)	Queens Plaza WB (E&F)	634	36.0	121.0
PATH (Newark)	Journal Square WB	478	37.0	204.0
Muni (San Francisco)	Montgomery WB (LRT)	2,748	32.0	129.0
SkyTrain (Vancouver)	Broadway EB	257	30.0	166.0
SkyTrain (Vancouver)	Metrotown EB (off-peak)	263	34.0	271.5
TTC (Toronto)	King SB	1,602	27.5	129.5
TTC (Toronto)	Bloor NB	4,907	44.0	135.0

Source: *TCRP Report 13* (*1*).
Note: NB: northbound, SB: southbound, EB: eastbound, WB: westbound

Method 3: Using Dwells from the Same System

This method only applies where a line of the same mode is being added to an existing system, in which case the controlling dwell time from an existing, similar, peak-point station can be used. Where passenger volumes at the critical station of the new line are different from the equivalent station on an existing line, the flow component of dwell time can be adjusted in proportion to hourly passenger movements in the station. Alternatively, the dwell time from an existing station with similar passenger volumes can be used.

Care should be taken if the train control system or operating procedures are different. If this is the case, consideration should be given to adjusting both the station dwell time and the operating margin.

Method 4: Calculating Dwells from Passenger Flows

TCRP Report 13 (*1*) provides regression equations to relate passenger flow times to the number of boarding, alighting, or mixed-flow passengers, and, in turn, to convert this flow time to dwell time. These regression equations can be used to estimate the dwell time from hourly passenger flows into the maximum load point station. However, the best regression fit involves logarithmic functions and the estimation of a constant for the ratio between the highest doorway and the average doorway passenger flow rate. The mathematics are complex and it is uncertain if the results provide any additional accuracy that merits this complexity—particularly if the hourly station passenger volumes by direction are themselves somewhat uncertain.

This method is best suited to new lines in locations without rail transit and with a sufficiently refined and calibrated regional transportation model that can assign hourly passenger flow, by direction, to individual stations. This method is not detailed further in this manual.

Step 1d: Select an Operating Margin

Section 4, Train Operations, introduced the need to add an operating margin to the minimum train separation and dwell time to create the closest sustainable headway without interference.

The closer the trains operate, and the busier they are, the more chance there is of minor incidents delaying service due to an extended station dwell time, stuck door, or late train ahead. It is never possible to ensure that delays do not create interference between trains, nor is there any stated test of reasonableness for a specific operating margin. A very small number of rail transit lines in the United States and Canada operate at capacity and so can accommodate little or no operating margin. On such lines, operations planners face the dilemma of scheduling too few trains to meet the demand, resulting in extended station dwell times and erratic service, or adding trains to the point that they interfere with one another. Striking a balance is difficult and the tendency in practice is to strive to meet demand—equipment availability and operating budget permitting. While the absolutely highest capacity is so obtained, it is poor planning to omit such an allowance for new systems.

The greater the operating margin that can be incorporated in the headway the better; systems running at design capacity have little leeway. The recommended procedure is to aim for 25 s and back down to 20 or even to 15 s if necessary to provide sufficient service to meet the estimated demand. Where demand is unknown or uncertain in the long-term future—when a rail line in planning reaches capacity—then 25 s should be used.

When a line already exists, the operating margin can be selected to accommodate 95% of dwells that occur at the critical station. This value can be estimated as twice the standard deviation of dwell times during the analysis hour (*1*).

When level boarding is not provided (i.e., high platform to high-floor car or low platform to low-floor car), the time required to operate to operate wheelchair lifts or deploy bridgeplates (discussed previously in Section 4) can be incorporated into the operating margin when the goal of the analysis is to determine the minimum headway that can be operated without causing train interference.

Step 1e: Determine the Non-Interference Headway

The non-interference headway is the sum of the train control separation determined in Step 1b, the average dwell time at the critical station determined in Step 1c, and the operating margin selected in Step 1d:

$$h_{ni} = t_{cs} + t_{d,crit} + t_{om}$$

Equation 8-4

where

h_{ni} = non-interference headway (s),

t_{cs} = train control separation (s),

$t_{d,crit}$ = average dwell time at the controlling station (s), and

t_{om} = operating margin (s).

Step 2: Determine the Minimum Headway Associated with the Right-of-Way Type

In some cases—most often with light rail transit lines—characteristics of the right-of-the-way may constrain the minimum headway that can be operated. A rail line that is completely grade separated and double tracked can skip this step. However, most light rail transit lines use a combination of right-of-way types. This step addresses three specific right-of-way types that create capacity constraints. These types are discussed in order of their decreasing relative importance for most systems. This order is as follows:

- Single track with two-way operation,
- On-street operation with traffic signal control, and
- Station departures adjacent to grade crossings.

Only the right-of-way types in this list that already exist or that are being evaluated need to be addressed in this step.

Step 2a: Single-Track, Two-Way Operation

Single track reduces capital costs but can add a serious capacity constraint.

This constraint only applies to two-way operation, not to one-way operation, such as on a downtown one-way street grid.

Single-track sections with two-way operation are the greatest capacity constraint on rail transit lines where they are used extensively. Single-track sections are used primarily to reduce construction costs. Some lines have been built with single track as a cost-saving measure where the right-of-way would permit double track. In other areas single track has been built because widening the right-of-way and structures is cost prohibitive. Single-track sections can be very short in order to bypass a particular obstacle, for example, an overpass of a highway.

While determining the potential extent of single-track construction is possible, the exact layout is highly system-specific. Estimates can be made of the number of track miles or kilometers required for a certain number of route miles or kilometers once the intended headway is known (*9*). While this does not tell the user *where* the single-track sections can be used, it can provide assistance in determining the possible extent of single track for use in cost estimates.

Gauntlet track interlaces the four rails without using switches.

Single-track sections greater than 0.25 to 0.30 mi (400 to 500 m) are a particularly restrictive capacity constraint. The minimum headway associated with single track is twice the time for trains to traverse the single-track section, including an allowance for switch throw and lock—unnecessary for spring switches or gauntlet track—and an operating margin to minimize the potential wait of a train in the opposite direction.

The time to cover a single-track section is:

Equation 8-5

$$t_{st} = S_m\left[\left(\frac{N_{st}+1}{2}\right)\left(\frac{3v_{max,st}}{d} + t_{jl} + t_{br}\right) + \frac{L_{st}+L_t}{v_{max,st}}\right] + N_{st}t_d + t_{sw} + t_{om}$$

where

t_{st} = time to cover single-track section (s),

L_{st} = length of single-track section (ft, m),

L_t = train length (ft, m),

N_{st} = number of stations on the single-track section,

t_d = average station dwell time (s),

$v_{max,st}$ = maximum speed on single-track section (ft/s, m/s),

d = deceleration rate (ft/s^2, m/s^2),

t_{jl} = jerk limitation time (s),

t_{br} = operator and braking system reaction time (s),

S_m = speed margin,

t_{sw} = switch throw and lock time (s), and

t_{om} = operating margin (s).

The speed margin is an allowance for out-of-specification equipment and train operators that do not drive at exactly the maximum permitted speed. It typically ranges from 1.08 to 1.20.

The minimum single-track headway is:

$$h_{st} = 2t_{st}$$

Equation 8-6

where

t_{st} = time to cover single-track section (s), and

h_{st} = minimum single-track headway (s).

Default values for light rail transit for use with Equation 8-5 are given in Exhibit 8-38. The results of applying these default values are depicted in Exhibit 8-39. A spreadsheet that calculates the single-track capacity is included on the accompanying CD-ROM.

A spreadsheet for calculating single-track capacity is included on the accompanying CD-ROM.

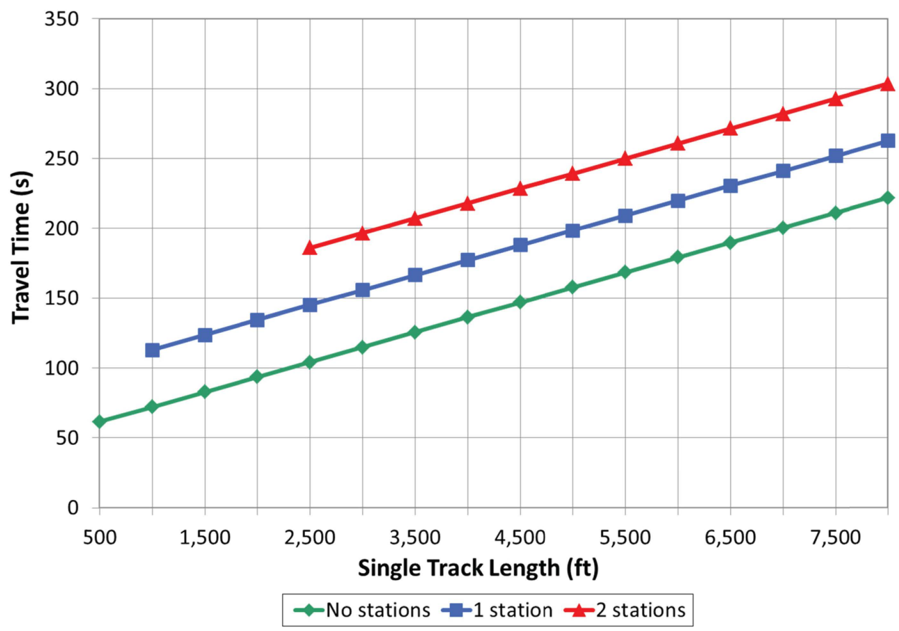
Exhibit 8-38
Default Data Values for Single Track LRT Travel Time

Term	Symbol	Value
Jerk limitation time	t_{jl}	0.5 s
Brake system reaction time	t_{br}	1.5 s
Dwell time	t_d	15–25 s
Switch throw-and-lock time	t_{sw}	6 s
Service braking rate	d	4.3 ft/s^2 (1.3 m/s^2)
Speed margin	S_m	1.1 to 1.2
Operating margin time	t_{om}	10–30 s

Source: *TCRP Report 13* (*1*).

Exhibit 8-39
Light Rail Travel Time Over Single-Track Section

The minimum single-track headway is twice the time given in this exhibit.



Source: *TCRP Report 13* (*1*).
Notes: Assumes speed limit of 35 mi/h, 180-ft train length, 20-s dwell time, 20-s operating margin, and other data as per Exhibit 8-38. *The recommended closest headway is twice the time given in the exhibit.*
A metric version of this exhibit appears in Appendix A.

The value used for the maximum single-track section speed should be the appropriate speed limit for that section. A speed of 35 mi/h (55 km/h) is a suitable value for most protected, grade-separated light rail lines. If the single-track section is on-street, then a speed below the traffic speed limit should be used. If there are signalized intersections, an allowance of half the signal cycle should be added to the travel time for each such intersection, adjusted for any improvements possible from traffic signal priority.

Scheduling for single track.

Trains should be scheduled from their termini such that passing locations are not close to the single-track sections. Where there is more than one single-track section, this can become difficult but not impossible.

Passing sections.

Lengthy single-track sections can severely limit headways and capacity and may require one or more double-track passing sections within the single-track section. These should, wherever possible, be of sufficient length to allow opposing trains to *pass on-the-fly* and allow some margin for off-schedule trains. Obviously, trains should be scheduled to meet at these locations.

Step 2b: On-street Operation

Reserved lanes for light rail vehicles and streetcars.

Historically, streetcar operation has achieved throughput in excess of 125 cars per hour on a single track in many North American locations, assuming relatively short one-car trains and line-of-sight operations. Some high-density systems continue to operate, such as the Queen Street Line in Toronto, where the Toronto Transit Commission operates single and articulated streetcars from multiple lines over a stretch of downtown streets at close headways. The price of this relatively high throughput capacity, however, is a low average operating speed, congestion, irregular running, and potential passenger confusion at multiple-car stops.

Despite this record, on-street operation is often raised as a major capacity constraint for modern light rail systems, yet this is rarely the case on contemporary lines. This is particularly true on most newer lines, where light rail trains have exclusive use of road lanes or a reserved center median where they are not delayed by other traffic making turns, queuing at signals, or otherwise blocking the path of the trains. Exclusive lanes for light rail are also being instituted on some of the older streetcar systems.

The capacity of streetcars operating in mixed traffic is determined similarly to the capacity of buses.

Single streetcars in classic mixed-traffic operation can be treated as similar to buses with capacity determined from the procedures of Chapter 6, Bus Transit Capacity, with suitable modifications reflecting longer vehicle lengths and differences in dwell time variability.

Where, as is often the case, light rail train lengths approach the downtown street block lengths, then the maximum train throughput is simply one train per traffic signal cycle, provided the track area is restricted from other traffic. When other traffic, such as queuing left-turning vehicles, prevents a train from occupying a full block, throughput drops as not every train can proceed upon receiving a green indication at a traffic signal. Similarly, longer-than-normal dwell times can cause a train to be unable to proceed on green, requiring the train to wait an additional traffic signal cycle to proceed. Therefore, a common rule of thumb is that the minimum sustainable headway is double the longest traffic signal cycle on the on-street portions of the line.

Equation 8-7 can be used to determine the minimum headway between trains operating on-street in exclusive lanes or mixed traffic (*1, 10*).

$$h_{os} = max\begin{cases} \dfrac{t_c + (g/C)t_d + Zc_v t_d}{(g/C)} \\ 2C_{max} \end{cases}$$

Equation 8-7

where

h_{os} = minimum on-street section train headway (s);

g = effective traffic signal green time (s);

C = traffic signal cycle length (s) at the stop with the longest dwell time;

C_{max} = longest traffic signal cycle length in the line's on-street section (s);

t_d = average dwell time (s) at the critical stop;

t_c = clearance time between trains (s), defined as the sum of the minimum clear spacing between trains (typically 15–20 s or the signal cycle time) and the time for the cars of a train to clear a station (typically 5 s/car);

Z = standard normal variable corresponding to a desired failure rate, from Exhibit 6-56 (page 6-65); and

c_v = coefficient of variation of dwell times (typically 40% for light rail, 60% for streetcars).

Some transit agencies use the signal cycle time (C) as the minimum clearance time.

Step 2c: Station Departures Adjacent to Grade Crossings

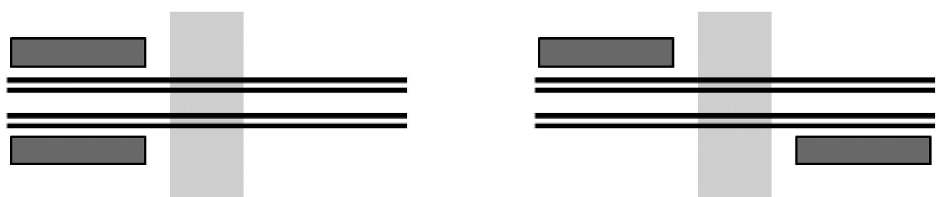
Grade crossings adjacent to stations.

Grade crossing activation and occupancy times can be affected by the presence of a station adjacent to the crossing. If the train must use the crossing after stopping at a station, the activation of the crossing signals is often premature and the crossing is unavailable to other traffic for more than the optimum time. In this case the train is also starting from a stop and so must accelerate through the crossing, adding to the total delay. Where the station platform is on the far-side of the crossing, the arrival time at the crossing can be predicted consistently and premature activation of the crossing is not a factor. The train is also either coasting or braking through the crossing from cruising speed and so will occupy it for less time.

Far-side platforms can be advantageous.

Stations can be designed to place both platforms on one side of the crossing or to locate one platform on each side of the crossing such that trains use the crossing before stopping at the station. Both arrangements are shown in Exhibit 8-40. Using far-side platforms is advantageous for the operational reasons given above, for reduced right-of-way requirements, and, for median operation, allowing left-turn bays to be readily incorporated into the street.

Exhibit 8-40
Light Rail Platform Options at a Crossing



(a) Facing (b) Far-side

Source: *TCRP Report 13* (*1*).

Additional time may need to be allowed when grade crossings adjacent to station exits are manually activated.

Delays to other roadway users caused by premature activation of crossing gates and signals at near-side stations can be reduced using wayside communication equipment. This can be done with the operator being equipped with a control to manually start the crossing cycle before leaving the station (as in Portland) or by an automatic method. An extra 10 s is an appropriate allowance when station departures are adjacent to grade crossings, and train operators manually initiate the crossing cycle after passenger movements at the station have finished. If the extra dwell time is required at a station with a dwell time equal or close to the critical dwell time along the line, then this station may end up controlling the headway:

Equation 8-8

$$h_{gc} = t_{cs} + t_{d,maxgc} + t_{gca} + t_{om}$$

where

h_{gc} = non-interference headway associated with stations with grade crossings on departure (s),

t_{cs} = train control separation (s),

$t_{d,maxgc}$ = longest average dwell time of stations with grade crossings on departure (s),

t_{gca} = minimum time from when the crossing cycle is manually activated to when a train can depart (s), and

t_{om} = operating margin (s).

An example of the automatic approach can be found on the San Diego Trolley. The trolley shares some of its track with freight trains and uses a communication device that identifies light rail trains to crossing circuits located on the far side of stations. If the crossing controller identifies a train as a light rail train, a delay to allow for station dwell time is added before the crossing is activated. This ensures that the crossing remains open for cross traffic for most of the time that the light rail train is stopped in the station. If the controller cannot identify the train as a light rail train, it assumes the train is a freight and activates the crossing gates without delay.

Other systems use an inductive link between the light rail train and wayside to activate signal preemption, switches, and ADA-mandated information requirements. The lowest-cost detection approach is the classic overhead contactor. Trolleybus technology using radio signals from the power collection pick-up to coils suspended on the overhead wires is also applicable to light rail but is not used in the United States or Canada.

Step 2d: Determine the Minimum Headway Associated with the Right-of-Way Type

The minimum headway associated with the right-of-way type h_{row} is the highest of the headways calculated in Steps 2a, 2b, and/or 2c:

Equation 8-9

$$h_{row} = \max(h_{st}, h_{os}, h_{gc})$$

where h_{row} is expressed in seconds and all other variables are as previously defined.

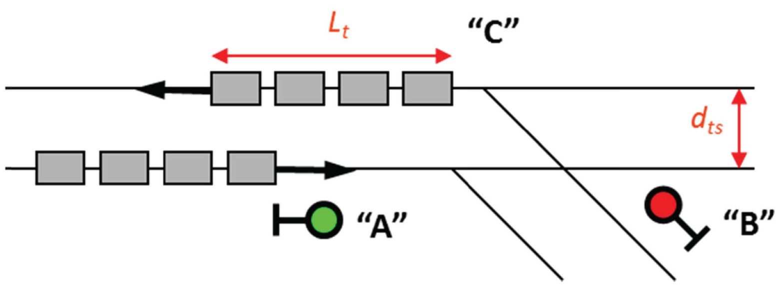
Step 3: Determine the Limiting Junction Headway

Junctions should not constrain capacity.

Correctly designed junctions should not be a constraint on capacity. Where a system is expected to operate at close headways, high-use junctions perform more reliably and at higher levels of capacity if they are grade separated. At such *flying junctions*, the merging and diverging movements can all be made without conflict and the only impact on capacity is the addition of the switch throw and lock times, typically 3 to 6 s. Speed limits, imposed in accordance with the radius of curvature and any superelevation, may reduce the schedule speed but should not raise the minimum headway—unless there is a tight curve close to a headway limiting station.

Key dimensions of a flat junction are shown in Exhibit 8-41.

Exhibit 8-41
Flat Junction Dimensions



Source: *TCRP Report 13* (*1*).

The worst case is based on a train (lower left) held at signal "A" while a train of length L_t moves from signal "B" to clear the interlocking at "C." The minimum operable headway is the train control separation of train "A" (imposed by the line's signaling system), plus the time required for the conflicting train to clear the interlocking, plus the extra time for train "A" to brake to a stop and accelerate back to line speed. Ignoring specific block locations and transition spirals, this can be expressed approximately as:

Equation 8-10

$$h_j = t_{cs} + \sqrt{\frac{2(L_t + 2f_{sa}d_{ts})}{a}} + \frac{v_{max}}{a+d} + t_s + t_{om}$$

where *[typical heavy rail values shown in brackets]*

h_j = limiting headway at junction (s);

t_{cs} = train control separation time (s); *[32 s]*

L_t = train length (ft, m); *[650 ft, 200 m]*

d_{ts} = track separation (ft, m); *[33 ft, 10 m]*

Higher-speed turnouts have smaller angles between the diverging tracks. They require a longer distance for the tracks to separate from each other, but trains can move onto the branch at a higher speed.

f_{sa} = switch angle factor (see also Exhibit 8-43):
—5.77 for #6 turnout,
—6.41 for #8 turnout, and
—9.62 for #10 turnout;

a = initial service acceleration rate (ft/s², m/s²); *[4.3 ft/s², 1.3 m/s²]*

d = service deceleration rate (ft/s², m/s²); *[4.3 ft/s², 1.3 m/s²]*

v_{max} = maximum line speed (mi/h, km/h); *[60 mi/h = 91 ft/s, 100 km/h = 27.8 m/s]*

t_s = switch throw and lock time (s); and *[6 s]*

t_{om} = operating margin time (s).

Although 120-s headways are possible, junctions generally should be grade separated for headways below 150–180 s.

Substituting the typical values shown above into the equation results in a junction limiting headway of 102 s. An operating margin should then be added to this headway. While in theory a flat junction should allow a 120-s headway, it does not leave a significant operating margin and there is a probability of interference headways. General guidance in rail transit design is that junctions should be grade separated for headways below 150 to 180 s.

Advantage of sophisticated supervision to reduce junction conflicts.

An exception is with a moving-block signaling system incorporating an automatic train supervision system with the capability to look forward. This system adjusts train performance and station dwells to avoid conflicts at the junction. That is, trains will not have to stop or slow down at the junction except for the interlocking's track design speed limit. In this case, the junction interference headway drops to 63 s, allowing 120-s, or slightly lower, headways to be sustained on a flat junction—a potentially significant cost savings associated with a moving-block signaling system.

Step 4: Check Power Supply Constraints

The power supply for a new rail line will presumably be designed to accommodate the number of trains planned to be operated, if not for the long term, then at least for some time into the future. However, the power supply for an existing line that is being considered for improved headways may not be capable of supporting the additional number of trains without being upgraded. The average headway imposed by a given substation will be a function of the number of trains the substation can power at a time and the time required for a train to traverse the track section powered by the substation, including station stops.

Step 5: Determine the Controlling Headway

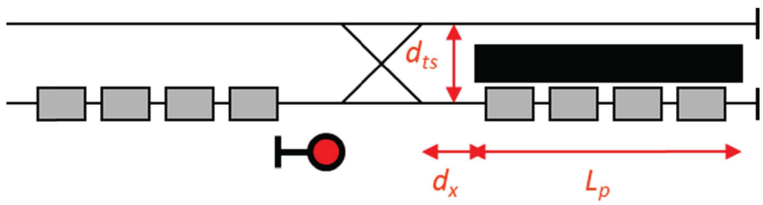
The controlling headway will be the highest of the non-interference headway h_{ni} (Step 1), the headway imposed by the right-of-way type h_{row} (Step 2), the highest limiting headway at a junction h_j (Step 3), and the minimum headway supported by the power supply system (Step 4). Assuming that turnbacks are not a capacity constraint (checked in Step 6), the controlling headway will be the minimum headway that can be consistently operated on the line.

Step 6: Determine Terminal Layover Time

Turnbacks should not be a constraint on capacity.

Correctly designed and operated turnbacks should not be a constraint on capacity. Key dimensions of a typical terminal station arrangement with a center (island) platform (preferred as passengers do not need to be directed to a particular platform; only a particular side of the same platform) are shown in Exhibit 8-42.

Exhibit 8-42
Key Turnback Dimensions



Source: *TCRP Report 13* (*1*).

The worst case is based on the arriving train (lower left) being held at the cross-over approach signal while a train departs. It must, moving from a stop, traverse the crossover and be fully berthed in the station before the next exiting train (lower right) can leave. The exiting train must then clear the crossover and the interlockings must be reset before another train can enter the station. The difference between the scheduled headway and the time required to make these maneuvers, doubled for a two-berth station such as the one illustrated, is available for terminal layover. The terminal layover time must be sufficient to accommodate passenger movements, and allow time for the driver to change ends, inspect the train, and check train integrity and braking. The maximum time available per track for terminal layover is given by Equation 8-11.

Equation 8-11

$$t_{tl} \leq 2\left(h - t_s - \sqrt{\frac{2(L_p + d_x + f_{sa}d_{ts})}{a+d}} - \sqrt{\frac{L_p + d_x + f_{sa}d_{ts}}{2a}}\right)$$

where *[typical heavy rail values shown in brackets]*

t_{tl} = terminal layover time (s);

h = train headway (s); *[120 s]*

t_s = switch throw and lock time (s); *[6 s]*

L_p = platform length (ft, m); *[660 ft, 200 m]*

d_x = distance from cross-over to platform (ft, m); *[65 ft, 20 m]*

d_{ts} = track separation (ft, m),

= platform width + 5.25 ft (1.6 m); *[33 ft, 10 m]*

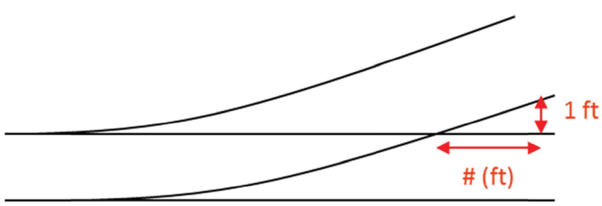
f_{sa} = switch angle factor (see also Exhibit 8-43):
—5.77 for #6 turnout,
—6.41 for #8 turnout, and
—9.62 for #10 turnout;

a = initial service acceleration rate (ft/s², m/s²); and *[4.3 ft/s², 1.3 m/s²]*

d = service deceleration rate (ft/s², m/s²). *[4.3 ft/s², 1.3 m/s²]*

Turnout numbers are based on the ratio of the number of feet moved longitudinally per foot moved out at the "frog" (device located where the tracks cross). The higher the number, the smaller the departure angle and the higher the permitted speed.

Exhibit 8-43
Turnout Numbers



Source: Wilson (*11*).

Terminal layover time can be calculated using the typical parameters given in the brackets above, including a headway of 120 s. The terminal layover time t_{tl} is less than or equal to 175 s per track. This value would increase by 9 s if the incoming train did not stop before traversing the crossover. While this is not a generous amount of time, particularly to contain a schedule recovery allowance, many systems maintain such close headways with minimal delays.

This analysis assumes that any speed restrictions in the terminal approach and exit are below the speed a train would reach in the calculated movements—approximately

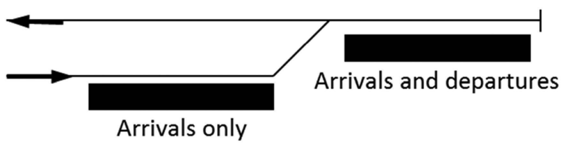
21 mi/h (34 km/h) on a stop-to-stop approach, and 29 mi/h (47 km/h) as the end of the train leaves the interlocking on exit. Normally there would be no restrictions so low, but following London Transport's Moorgate disaster—when a fully loaded train accelerated into the wall at the end of a terminal station—some systems have imposed low entry speeds, occasionally enforced with speed control signaling.

The maximum permitted terminal time can be calculated for the specific system and terminal parameters, using the controlling headway as the value for *h* in Equation 8-11. This time should then be compared to the time required to load and unload passengers, perform necessary checks, and have the driver walk to the opposite end of the train. There are numerous corrective possibilities when the terminal time is insufficient. These include moving the crossover as close to the platform as possible; however, structures can restrict the crossover location in subways.

Toronto's streetcars face terminal design problems in which two or more routes share a common terminal and single-track turning loop. This is the case at the Broadview and Dundas West subway stations where there is heavy transferring activity between the subway and streetcars. The high volumes of transit vehicles and passengers can cause delays to the following streetcars while passengers board and alight from the preceding car. Scheduled recovery time for the streetcar operator is hard to accommodate in these conditions as the volume of the following cars precludes layover time.

The Baltimore light rail line also uses single-track termini but the train frequency is not high enough for these to be a capacity limitation. However, some terminals are designed to allow an arriving train to unload passengers before the departing train ahead leaves. This is accomplished through the use of an extra platform as shown in Exhibit 8-44. This arrangement allows the location of a station in a relatively narrow right-of-way since the platforms are not adjacent to each other and a wider center platform is not required.

Exhibit 8-44
Light Rail Single-Track Terminus with Separate Unloading Platform



Source: *TCRP Report 13* (*1*).

Dual-faced platforms and loops can reduce dwell times.

If passenger dwell is a limiting factor, then this issue can be reduced with the use of dual-faced platforms. At terminals with exceptionally heavy passenger loading, multiple-track layouts may be needed. Another alternative, used at SEPTA's 69th Street, New York's South Ferry, and Chicago's Howard and Forest Park termini, are loops—however, these are rare luxuries for heavy rail transit. However, some older streetcar-based light rail lines still incorporate terminal loops.

At a leisurely walking pace of 3 ft/s (1 m/s), it takes 200 s for an operator to walk the length of a 650-ft (200-m) heavy rail train, more if the operator is expected to check the interior of each car for left-behind objects or passengers. Obviously, this cannot be accommodated reliably in a 175-s terminal layover time. The turnaround time can be expedited with *fall-back* or *set-back crewing*, where an extra engineer or operator is positioned at the rear end of the terminal platform as a train enters the station. The fallback engineer enters the rear cab of the train as the train's incoming engineer exits the front cab. Once the train has been readied for departure and received its boarding passengers, it can depart without further delay. In this type of operation, each engineer drops back to the following train at the terminal station, so only one additional engineer per shift is required to make such an operation work. Such a strategy can improve equipment utilization and reduce the car requirements associated with special events, and can also be productive during normal peak periods at high-volume terminal stations.

Allowances should be made to prevent common delays from disrupting terminal operations.

Terminal arrangements should accommodate some common delays. An example would be the typical problems of a train held in a terminal for a door sticking problem, waiting for police to remove an intoxicated passenger, or for a cleaning crew to perform minor cleaning. Alternatively, one track may be preempted to store a disabled train. On these occasions, the terminal is temporarily restricted to a single track and the maximum terminal layover time is reduced to 61 s with the above parameters (70 s without an approach stop). This may be sufficient for the passenger dwell but cannot accommodate changing ends on a long train and totally eliminates any schedule recovery allowance.

Alternative solutions for terminals include single-track terminals (which limit the overall throughput capacity), turnback loops, which are prevalent among many older streetcar systems, and terminal stations with far side crossovers and tail tracks (the arriving train pulls into one platform, then pulls into one of the tail tracks, changes direction and then returns to pick up passengers from the other platform).

More expensive ways to improve turnbacks include multi-track or grade-separated terminals, or extending tracks beyond the station and providing crossovers at both ends of the station. This permits a storage track or tracks for spare and disabled trains—a useful, if not essential, failure management facility. With crossovers at both ends of the station, on-time trains can turn beyond the station with late trains turning in front of the station—providing a valuable recovery time of about 90 s at the price of additional equipment to serve a given passenger demand.

If terminal time is insufficient and none of the corrective measures discussed above can be implemented, then the terminal becomes the capacity constraint and the value of h that produces a sufficient terminal time then becomes the controlling headway.

Step 7: Determine Train Throughput

The maximum number of trains per hour T (line capacity) is simply the number of seconds in an hour divided by the controlling headway determined from Step 5 (or Step 6, if for some reason a turnback constrains capacity):

$$T = \frac{3{,}600}{h_c}$$

Equation 8-12

where

T = line capacity (trains/h),

3,600 = number of seconds in an hour, and

h_c = controlling headway (s/train).

Step 8: Determine Person Capacity

The person capacity of a rail route at its maximum load section *under prevailing conditions* is determined by multiplying the number of cars operated during the peak hour by the agency's scheduled design load for each car and by a peak-hour factor that reflects loading diversity:

Equation 8-13

$$P = P_c C_h (PHF)$$

where

P = design person capacity (p/h),

P_c = maximum design load per car (p/car),

C_h = cars operated per hour (car/h), and

PHF = peak-hour factor.

Where an agency operates a mix of cars along the line (e.g., a mix of high-floor and low-floor cars), a weighted average design load should be used, based on the proportion of each car type used. If the transit agency has established a loading standard for its rail cars, this value should be used. Otherwise, the procedures given in the Passenger Load section of Chapter 5, Quality of Service Methods, may be used to estimate a maximum design load based on a car's exterior dimensions and interior seating arrangement.

The person capacity of a rail route at its maximum load section *when operated at line capacity* is determined by multiplying the number of trains per hour by the number of cars per train, the agency's scheduled design load for each car, and a peak-hour factor:

Equation 8-14

$$P = T N_c P_c (PHF)$$

where

P = design person capacity (p/h),

T = line capacity (trains/h),

N_c = number of cars per train (cars/train),

P_c = maximum schedule load per car (p/car), and

PHF = peak-hour factor.

The peak-hour factor reflects the diversity of demand over the course of a peak hour and produces a person capacity that reflects the number of people that can consistently be served day after day at the desired loading (quality of service). It is defined as:

Equation 8-15

$$PHF = \frac{P_h}{4P_{15}}$$

The PHF ranges from 0.25 (all volume occurs during the peak 15 min) to 1.00 (volumes are even throughout the hour).

When 30-min peak periods are used, P should represent 30-min volumes, and $2P_{30}$ should be substituted for $4P_{15}$.

where

PHF = peak-hour factor,

P_h = passenger volume during the peak hour (p), and

P_{15} = passenger volume during the peak 15 minutes (p).

PHFs observed at many U.S. and Canadian rail systems in the mid-1990s are tabulated in Exhibit 8-45, illustrating the variations in peak-hour loading diversities that can occur among rail transit systems. Use of current local data is recommended whenever possible; when these data are not available, the following default PHF values for specific modes can be used instead:

- 0.80 for heavy rail,
- 0.75 for light rail, and
- 0.60 for commuter rail operated by electric multiple-unit trains.

Exhibit 8-45
Observed Rail Transit Peak-Hour Factors (1994-95)

System (City)	# of Routes	Peak Hour Factor
Commuter Rail		
AMT (Montréal)	2	0.71
CalTrain (San Francisco)*	1	0.64
GO Transit (Toronto)*	7	0.49
Long Island Rail Road (New York)	13	0.56
MARC (Baltimore)*	3	0.60
MBTA (Boston)*	9	0.53
Metra (Chicago)	11	0.63
Metro-North (New York)	4	0.75
NICTD (Chicago)	1	0.46
New Jersey Transit*	9	0.57
SCRRA (Los Angeles)*	5	0.44
SEPTA (Philadelphia)	7	0.57
VRE (Washington, D.C.)*	2	0.35
Light Rail		
CTS (Calgary)	2	0.62
RTD (Denver)	1	0.75
SEPTA (Philadelphia)	8	0.75
TriMet (Portland)	1	0.80
Rapid Transit		
SkyTrain (Vancouver)	1	0.84
CTA (Chicago)	7	0.81
MARTA (Atlanta)	2	0.76
Metrorail (Miami)	1	0.63
NYCT (New York)	23	0.81
PATH (New Jersey)	4	0.79
STM (Montréal)	4	0.71
TTC (Toronto)	3	0.79

Source: *TCRP Report 13* (*1*).
Note: *Mainly diesel-hauled—not electric multiple unit.

When the specific vehicle type has not yet been selected (e.g., when planning a new rail system), vehicle length can be used as a proxy for the passenger capacity of a rail car. Passenger loadings for typical North American light rail cars range from 1.5 to 2.4 passengers per foot of car length (5.0 to 8.0 p/m length). The lower level of 1.5 passengers per foot length (5.0 p/m length)—with a standing space per passenger of 4.3 ft^2 (0.4 m^2)—corresponds to a standing load without body contact, while the upper level provides 3.2 ft^2 (0.3 m^2), corresponding to a standing load with some body contact.

Chapter 5, Quality of Service Methods, describes passenger conditions at different loading levels.

For heavy rail, the 75-ft (23-m) cars used in more than 12 U.S. and Canadian cities range from 2.1 to 3.5 passengers per foot of car length (7.0 to 11.5 p/m of car length). The higher end of this range approaches crush-loaded conditions. The lower end of the range, at 2.1 to 2.4 passengers per foot length (7 to 8 p/m length)—with a standing space per passenger as low as 3.2 ft² (0.3 m²)—is an appropriate and tight range for higher use systems. A lower figure of 1.8 p/ft length (6 p/m length) provides a more comfortable loading level and is appropriate for a higher quality of service on new systems. In either case, a reduction by 0.3 p/ft length (1.0 p/m length) should be used for smaller, narrower cars (*1*). Actual passenger loading standards in use at some agencies may differ from these values and should be used for analyses related to those systems.

COMMUTER RAIL CAPACITY

Overview

Commuter rail ridership in North America is dominated by the systems in the New York area where the busiest routes use electric multiple-unit trains on dedicated tracks with little or no freight service. The capacity of such systems can be estimated using the generalized rail transit methodology presented above, with suitable adjustments to input parameters to account for the sometimes lower vehicle performance and lower throughput of signaling systems where these are based on railroad rather than rapid transit practices. These high levels of throughput generally are limited to the commuter rail systems feeding Grand Central Terminal and Penn Station in New York City. Other commuter rail systems, such as those feeding Chicago, Washington, D.C., Boston, Philadelphia, Los Angeles, and San Francisco, operate at relatively short headways into and out of the major terminals—but at a level of peak-hour throughput per track significantly below the New York systems.

Push-pull commuter rail operation.

Outside of New York, with the exception of SEPTA's Philadelphia lines, Chicago's Metra Electric and South Shore lines, and Montréal's Deux-Montagnes line, commuter rail uses diesel locomotive-hauled coaches and follows railroad practices. Most of these commuter systems operate in *push-pull* mode, where the locomotive always remains at the same end of the train, and a cab control car is positioned at the other end of the train. In *pull* mode, the engineer operates the train in the "normal" forward direction, from the locomotive at the front of the train. In *push* mode, the train engineer operates the train from the front of the cab car and the locomotive is at the rear end of the train. The top speed of commuter trains in push mode is limited by the crashworthiness of the cab car and can be capped at 100 mi/h (160 km/h) or, in some cases, slightly higher. Push-pull operations reduce train layover times, minimize operating costs, and increase the available capacity at terminals.

Electric locomotive-hauled coaches are also used by SEPTA and New Jersey Transit on routes that also run electric multiple-unit cars. Dual-powered (electric and diesel) locomotives are used by the Long Island Rail Road (LIRR) and Metro-North Railroad in the New York area. Most new starts are likely to use diesel locomotive-hauled coaches; however, three of Denver's new commuter rail lines (under construction at the time of writing) will use electric multiple-unit cars.

For non-electric lines, there is no easy answer for calculating capacity. Unlike rapid and light rail transit, whose vehicles have similar performance characteristics within their respective modes, the performance of diesel locomotives used by various U.S. commuter operations varies considerably. This performance, measured by the *power-to-weight (P/W) ratio,* ranges from 2 to 10 for commuter rail operations, which makes it difficult to develop a "standard" commuter rail locomotive performance for use in capacity calculations. For comparison, a typical diesel Amtrak intercity train has a P/W ratio of 4 to 6, while electric high-speed corridor trains (such as the Metroliner used in the Washington-New York-Boston corridor) have P/W ratios of 10 and higher (*12*). Exhibit 8-46 shows the effect of different P/W ratios on the time and distance needed to accelerate from a stop, and the delay incurred as a result.

Exhibit 8-46
Effect of P/W Ratio on Train Acceleration to 80 mi/h (128 km/h)

	Power-to-Weight (P/W) Ratio				
	2	**4**	**6**	**8**	**10**
Distance to accelerate (mi)	23.0	7.3	3.6	2.5	1.9
Distance to accelerate (km)	37.0	11.8	5.8	4.0	3.1
Acceleration time (min)	23.7	7.7	4.3	3.0	2.3
Time lost (min)	3.7	2.3	1.6	1.2	1.0

Source: Galloway (*12*).

Other issues affect commuter rail capacity that make it difficult to provide a simple analytical technique. First, many smaller commuter rail lines do not own the tracks they use, and therefore the number of trains they can operate will depend on negotiations with the owning railroad. Second, the mix of users of the tracks—and their impacts on capacity—will vary greatly from location to location. Generally, simulation is the only tool available for calculating the capacity of these commuter rail lines. Finally, the number of platforms available at terminal stations may constrain capacity. Consequently, this section does not present any equations for calculating commuter rail capacity. Instead, it focuses on the factors that impact capacity and potential means of improving capacity.

Simulation is often the only tool available for calculating commuter rail capacity.

Track Ownership and Usage

For commuter rail lines that use tracks owned by another railroad, the number of trains that can be operated in the peak hour depends on negotiations with the owning railroad. As the number of trains using a track increases—particularly when only a single track is available—the average speed of all trains decreases. Train meets have a compounding impact on capacity: each meet produces delay to the train that must wait, and each delay produces an increased probability of additional future meets. The impact of meets is even more severe when different classes of trains with different characteristics (e.g., passenger and freight trains) share the same tracks (*13*).

One concern that a freight railroad will have when passenger trains are proposed to be added to its tracks will be the impacts on train running times. Train crews have a maximum permitted number of hours they can work at a time, and an increase in train travel time may put them at risk of exceeding that limit if any unexpected delay occurs. Freight railroads may also need to reserve capacity (paths) for freight trains to service local customers during hours that passenger service is being contemplated, or to get trains to a certain location by a certain time (*14*).

There are a number of consumers of track capacity, some recurring but most not. The most common consumers of capacity are (*13*)

- Trains (not all use the same amount),
- Track patrols,
- Track maintenance,

- Track deterioration requiring temporary speed restrictions,
- Passenger station stops,
- Industrial switching,
- Freight yard interactions,
- Train or train control system failures,
- Incidents (e.g., crossing accidents, deer, and trespassers), and
- Weather.

Trains will be assigned different levels of priority, and there may be different levels of priority within a particular class of trains. For example, passenger train types can include high-speed intercity, conventional intercity, commuter zone express, commuter local, and deadhead (non-revenue) passenger trains. Freight train types include intermodal, manifest, bulk commodity, and local freight. An individual train's priority may also be raised or lowered depending on special circumstances. For example, early trains will have lower priority, trains whose crews are nearing their legal work hour limit will have higher priority, and heavy trains may be given higher priority, particularly on grades, because of the time required to regain their speed after a stop. The relative priority of each train will determine which one is delayed when two trains meet or one overtakes another (*14*).

Although freight railroads are becoming more receptive to accommodating commuter rail services—and the revenue and capital upgrading they produce—they have the upper hand and obtaining paths for commuter trains at a reasonable cost can require difficult and protracted negotiations.

Transit agency ownership of track used for commuter rail.

There are an increasing number of exceptions where the operating agency has purchased trackage and/or operating rights and so has more, or total, say in the operation and the priority of passengers over freight. The two New York commuter railroads own the great majority of track on which they operate; however, in the case of MTA-Metro North Railroad, priorities must be determined between Metro North's commuter operations and Amtrak's Northeast Corridor services. New Jersey Transit, SEPTA in Philadelphia, MBTA in Boston, Metra in Chicago, and Metrolink in Los Angeles, among others, also own substantial portions of the trackage they use. Some transit agencies have leverage with the freight railroads, as they own tracks used by the freight carriers. However, even when an agency owns track or trackage rights, there may still be strict limits on the number of trains that can be operated because of interlockings and grade crossings with other railroads.

Train Throughput

Determining train throughput requires consulting the railroad agreement or the railroad or agency signaling engineers to determine the maximum permitted number of commuter trains per hour. Generally these numbers will be based on a train of maximum length.

A definitive answer may not always be obtained, particularly with single-track sections that are shared with freight. Freight traffic can be seasonal and available commuter rail trips can vary. Usually the agreement will ensure a minimum number of commuter rail trips per hour. These may be unidirectional—that is, all trains must platoon in one direction in each peak period. This is generally not a capacity problem but rather an efficiency issue with respect to equipment and staff utilization. Unidirectional operation is an issue on lines where reverse commuting to suburban work sites is important. For example, Chicago's Metra has services aimed specifically at the growing reverse commuter market.

Signal blocks for freight trains are considerably longer than for rail transit operations, due to the length of the trains, and the amount of time and distance required to bring a long, heavy freight train to a stop. Trains are the only means of land transport that cannot stop within their range of vision (*15*). Because of these long stopping distances and the resulting longer block lengths, and the lower speed of freight trains compared with rail transit, both commuter and freight trains take longer to traverse a signal block than their rail transit counterparts. This longer block transit time translates into significantly longer headways between trains and, therefore, lower capacity.

Line Capacity Range

Train throughput where commuter rail has exclusive occupancy of the track.

The number of commuter rail trips available per hour may range from one to the double digits. Ten or more trains per hour is at the upper range of traditional railroad signaling and will exceed it if long, slow freights must be accommodated. At the upper end of this range, commuter rail is effectively in sole occupancy of the line for the peak period and is approaching levels where the generalized rail transit capacity methodology can be considered. The input values should be adjusted using suitably lower braking and acceleration rates and longer train lengths, and by adjusting the separation safety factor b from the suggested value of 2.4 for a rapid transit three-aspect signaling system to 3 or 4. This equation and the associated equation for junction throughput do not apply in locations and times where freight and commuter rail trains share trackage or where the signaling system is designed solely for freight with long signal blocks.

Operating practices and patterns.

Additional complications are raised by the variety of commuter services operated and the number of tracks available. The busier commuter rail lines tend to offer a substantial number of stopping patterns in order to minimize passenger travel times and maximize equipment utilization. A common practice is to divide the line into zones with trains serving the stations in a zone and then running express to the station(s) in the CBD. Through local trains provide connections between the zones. A number of lines in the Chicago and New York areas are operated this way—Metra's Burlington Northern line to Aurora operates with five zones in the morning peak, Metro-North's New Haven line (including the New Canaan Branch) operates with seven zones. Such operating practices are made possible with three or more tracks over much of the route and the generous provision of interlockings to allow switching between tracks. Grade-separated junctions are also common where busy lines cross or converge. The capacity of this type of operation is hard to generalize and should be considered on a case-by-case basis. Such heavy operations are similar to grade-separated rapid transit in many ways, but have some notable exceptions, such as the wide range of services operated.

Station Constraints

Train storage at downtown terminals.

Another principal difference between commuter rail and the other rail transit modes is that commuter rail trains are often stored at the downtown terminals during the day. This reduces the need for track capacity in the off-peak direction and allows a higher level of peak-direction service to be operated. Metro-North in New York, with 42 platform tracks in use at Grand Central Terminal, is thus able to use three of its four Park Avenue tunnel tracks in the peak direction. Even when one of the tunnel tracks was closed for reconstruction, 23 trains per hour were handled on the remaining two peak-direction tracks.

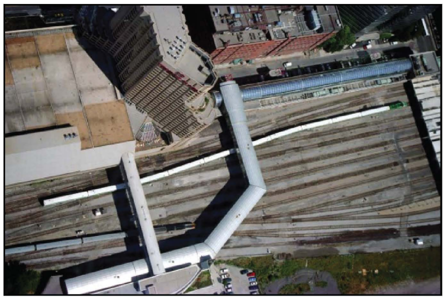
The situation at New York's Penn Station is less relaxed. The LIRR has exclusive use of five tracks and shares four more with Amtrak and New Jersey Transit. At the time of writing, the LIRR operated the East River tunnels with two tracks inbound and two tracks outbound, with an average peak headway of approximately 3 min per track. With limited station capacity, many LIRR trains continue beyond Penn Station to the West Side Yard during rush hours. However, the yard cannot accommodate the full complement of peak trains, and some trains must be turned in the station. This can be done in as little as 3.5 min in a rush, but a minimum of 8 min is generally scheduled for turning trains.

Station Dwells

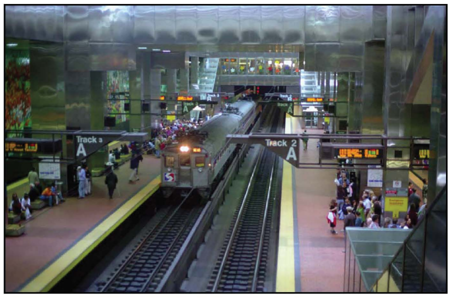
Dwells are less critical for commuter rail than for heavy rail transit.

Station dwell times on commuter rail lines are generally not as critical as they are on rapid transit and light rail lines, as frequencies are lower and major stations have multiple platforms, such as those shown in Exhibit 8-47. In most cases, the longest dwells are at the multiple-platform downtown terminals where the train is not blocking others while passenger activity takes place.

Exhibit 8-47
Multiple-Platform Commuter Rail Terminal Examples



(a) Toronto



(b) Philadelphia

Platform level and commuter rail car door layout.

Passenger flows are generally unidirectional and so are not slowed by passengers attempting to board while others alight and vice-versa. Exceptions are locations where major transferring activity takes place between trains but these are limited. Jamaica station on the LIRR is one of the few examples of a station with major transfers as it serves as a funnel where eight lines converge from the east and two major lines diverge to the west. Most transfers are made cross platform and are scheduled for 2 or 3 min. SEPTA's four-track regional rail tunnel through Center City Philadelphia—where train schedules incorporate both dwell and schedule recovery time—and the FrontRunner line in Salt Lake City are among the few North American locations where commuter trains run through from one line to another without terminating downtown.

Commuter rail station dwell times depend on the platform level and car door layout. The busiest lines are equipped with high platforms and remotely controlled sliding doors, as on rapid transit cars. Single-level cars often use conventional traps for high- and low-platform stations but these are time consuming to operate and require a large operating crew. Cars used on lines with both high and low platforms can be fitted with conventional trap doors at the car ends and sliding doors for high-platform use at the center of the car, as on New Jersey Transit, the South Shore in Chicago, and the Deux-Montagnes line in Montréal. Most bi-level and gallery cars are designed for low platforms and have the lowest step close to the platform for easy and rapid boarding and alighting. Bi-level cars of the type popularized by GO Transit feature two automatic sliding double stream doors per side allowing cars to be emptied in 1 to 2 min. Gallery cars usually feature one exceptionally wide door (6.5 ft or 2 m) at the center of each side to allow rapid boarding and alighting with multiple passenger streams.

Means of Increasing Line Capacity

If the line capacity is determined to be insufficient for the desired level of commuter rail operations, there are three main ways that capacity can be increased: (a) add another track, (b) reduce running times between sidings, and (c) reduce the delay resulting from train meets and overtakes (*1*).

Methods that do one or more of these things are described below, along with a qualitative discussion of each method's potential benefits and potential constraints. Simulation will be required to quantify the effects of a particular method for increasing capacity.

Double Tracking

Double tracking allows some trains to meet without having to stop. Double-track sections can be formed by joining or extending existing sidings, but need to be at least three signal blocks (4.5 to 7.5 mi or 7 to 12 km) long in order to be effective. Longer double-track sections should provide crossovers to allow both meets and overtakes to occur within the double-track section (*1*).

The ends of sidings or double-track sections should not be located on or near heavy grades (1% or more) because of the difficulty of starting and stopping heavy trains. Curves should also be avoided at the ends of double-track sections because of the difficulty of installing and maintaining switches located on curves. Finally, grade crossings should not be located near the ends of double-track sections because they would be blocked by a train stopped for a meet or overtake (*1*).

Avoid ending sidings and double-track sections

- *On or near grades,*
- *On curves, and*
- *Near grade crossings.*

At the extreme, the entire line can be double-tracked. Adding double track to all of Tri-Rail's line in South Florida allowed it to increase service from one passenger train per direction per hour to three. However, the cost of double-tracking a long rail line can be very high, particularly when bridges or tunnels are required, or when additional right-of-way must be acquired.

Double-tracking an entire line, while very beneficial from a capacity standpoint, is also very costly.

Adding and Lengthening Sidings

Shorter sections of double track are known as sidings. When trains meet at a siding, one will need to stop and wait for the opposing train to pass (sections of the line that are considered "double track" are long enough for some trains to meet without having to stop). Trains experience two types of delay at sidings, *fixed* and *variable.* Fixed delay includes delay associated with decelerating, stopping, and accelerating, as well as any difference in operating speed between the siding and the main line. Variable delay consists of time that a train must wait for the opposing train once it is in the siding (*11*).

Sidings are short sections of double track not long enough for trains to pass without one having to stop.

Increasing the number of sidings reduces variable delay, as trains can be directed to a siding closer to the time a meet will occur, but does not change the fixed delay associated with stopping. The capacity benefit diminishes as each new siding is added because the variable delay is reduced by smaller and smaller increments, but the fixed delay remains (*1*).

Lengthening sidings reduces variable delay, because the distance between the ends of sidings is reduced. However, it adds to fixed delay, because the amount of time required to transit a siding increases as the siding's length increases (*1*).

Providing Higher-Speed Siding Entries and Exits

Increasing the siding entry speed may also require improvements to the siding itself.

Fixed delay is reduced when trains can enter and exit the main line at higher speeds. A siding's entry and exit speed is controlled by the angle of departure of the siding from the main line, which is measured by the switch number (see Exhibit 8-43, page 8-62). The higher the switch number, the faster the entry and exit speeds. Sidings must permit speeds at least as high as the entry and exit speeds, must be signaled, and must be long enough to allow a train to stop from the higher entry speed (*1*).

Train Control System Improvements

Signals can be moved closer together, which shortens block lengths and permits trains to run closer together, within the limits created by the safe braking distance needed for the worst-case train. Changing the signal spacing mainly reduces delay when one train overtakes another, as the overtaken train can depart sooner once the other train has passed. Shortening the lengths of blocks can also create a minor improvement in meet delay, as dispatchers have better information about train positions to help them make decisions about at which siding to have trains meet (*1*).

Infrastructure Improvements

Improving track conditions to improve train speeds may not improve capacity if blocks have to be lengthened to accommodate faster trains.

Track conditions on a railroad being considered for commuter rail service may restrict trains' maximum speed. The Federal Railroad Administration defines various track classes, based on such factors as curvature, superelevation, track condition, number of crossties per unit length, and so forth, and sets maximum allowed passenger and freight train speeds based on those classes. Infrastructure improvements to upgrade the track class will improve train operating speeds; however, capacity may not change, as signal blocks may need to be lengthened to safely accommodate the higher speeds, resulting in little or no net change in time to transit a block. (In cases where no train signaling exists—*dark territory*—a signaling system will need to be developed.) Exhibit 8-48 shows the maximum speeds permitted for different track classes. Lower regulatory speeds may apply, depending on the type of train control system being used.

Exhibit 8-48
U.S. Railroad Track Classes

	Passenger		Freight	
Track Class	**mi/h**	**km/h**	**mi/h**	**km/h**
Excepted	Not allowed	Not allowed	10	16
1	15	24	10	16
2	30	48	25	40
3	60	96	40	64
4	80	128	60	96
5	90	144	80	128

Source: Code of Federal Regulations, Title 49, Part 213.
Note: Track classes 6 and higher, not shown, are used for high-speed intercity passenger rail.

Other infrastructure issues can create capacity constraints (*1*):

- *Junctions* are often under the control of different dispatchers, requiring a train to be held at a junction, blocking the exit. Providing a siding at the junction can mitigate this problem.
- Trains operate more slowly when entering and exiting *freight yards.* Providing the ability to bunch trains, either through closer signal spacing or additional tracks can mitigate impacts on capacity. Older freight yards may have been designed for shorter trains; the yard entry track needs to be sufficiently long to hold an entire train without blocking the mainline, while yard switches are lined manually.
- Cars may be temporarily stored on the mainline during *switching operations* on industrial tracks. A service track to store these cars can be constructed to mitigate this problem.

Commuter Rail Operating Speeds

Exhibit 8-49 gives average commuter rail operating speeds, including station stops, for different combinations of P/W ratios, station spacings, and dwell times. The exhibit assumes a combination of conventional block signaling and track conditions providing a passenger train speed limit of 79 mi/h (127 km/h), no grades, and no delays due to other trains. Note that in most cases, except for the higher P/W ratios and longer station spacings, a train will not be able to accelerate to the assumed speed limit before it has to slow for the next station. When the characteristics of the line (e.g., grades and station locations) and equipment to be used are known, a train simulator should be used to estimate operating speeds. A dwell time of 30 s would be difficult to achieve on a higher-volume line, but might be appropriate for lower-volume lines and off-peak periods (*12*).

Exhibit 8-49
Average Commuter Rail Operating Speeds

	Average Operating Speed (mi/h)		
Station Spacing (mi)	**P/W = 3.0**	**P/W = 5.8**	**P/W = 9.1**
Average Dwell Time = 30 s			
1.0	16.8	20.3	22.3
2.0	25.8	30.9	35.0
4.0	36.4	44.1	48.6
5.0	40.3	48.7	52.7
Average Dwell Time = 60 s			
1.0	14.8	17.4	18.8
2.0	23.3	27.4	30.6
4.0	33.8	40.4	44.1
5.0	37.8	45.0	48.5

Source: Galloway (*12*).
Notes: P/W = power-to-weight ratio. Values assume 79 mi/h speed limit, no grades, and no delays due to other trains. A metric version of this exhibit appears in Appendix A.

Person Capacity

Except for a few situations in which standing passengers are accepted for short distances into the city center, commuter rail person capacity is based solely on the number of seats provided on each train. A peak-hour factor is used in the rail capacity methodology to develop a design hourly capacity that allows for variations in passenger boarding demand between trains during the peak hour. For individual trains, a person capacity based on 90% of a seated load is a reasonable value that reserves capacity to accommodate higher-than-average passenger demands on a given trip.

Car seating capacities vary by car type and by rail operator. Generally, single-level coaches or electric multiple-unit cars have seating capacities in the range of 80–90 for a two-by-two seating configuration, and 110–116 for a three-by-two configuration. Bi-level coaches are becoming more prevalent among U.S. commuter rail systems. These come in multiple configurations but generally offer two-by-two seating, with seating capacities for the newest models ranging from 135 to 150 passengers per car.

Maximum train lengths (or consists) are another determining factor in the person-carrying capacity of a train. The maximum train length will be governed by factors unique to each commuter rail system, including:

- Station platform lengths (which can be limited by local conditions such as right-of-way, horizontal and vertical track curvature, and the proximity of turnouts or track switches);
- Yard track lengths;
- The maximum number of cars that can be hauled by the model of locomotive being used for the service; and
- Availability of rail cars to lengthen trains.

AUTOMATED GUIDEWAY TRANSIT CAPACITY

AGT has nuances that must be considered when applying the generalized rail transit capacity methodology.

In general, AGT capacity can be estimated using the generalized rail transit capacity method presented earlier in this section. However, there are some nuances specific to AGT that must be considered that are discussed below.

AGT Technology

AGT provides a very small share of urban, public, fixed-guideway transit—being used for less than 0.1% of passenger trips in the United States—but its use increases when institutional systems are considered, most of which are intra-airport shuttles. Technology ranges widely from standard-gauge advanced light rapid transit (e.g., Vancouver’s SkyTrain), to the downtown people-mover in Miami, to small-scale monorails in amusement parks. All AGT systems are proprietary designs. As such, AGT vehicle performance, acceleration, and braking rates vary greatly, as does their balance between speed, vehicle size, and capacity.

Train Control Separation

Train control systems on AGT range from sophisticated moving-block signaling systems to basic manual systems in which only one train may be on a section of line—or the entire line—at a time. Manual or radio dispatching may ensure that a train does not leave a station until the leading train has left the station ahead. One variation uses sectioned power supply. Power is disconnected for a given distance behind an operating train.

These operating variations are not fully accommodated in the generalized method. However, if the basic AGT performance criteria are known, then the method will provide an approximation of the minimum train separation time for a range of AGT train controls—from a moving-block signaling system to a simple fixed-block system. A surrogate of this can be roughly simulated by setting the train detection uncertainty factor B at four times the minimum braking distance. The results are shown in Exhibit 8-50 for trains of typical AGT lengths, using the specific AGT values in Exhibit 8-51, with terms adjusted from typical rail transit values shown in bold.

Exhibit 8-50
AGT Minimum Train Separation Times

Train Length	Minimum Train Separation (s) Fixed Block	Moving Block
160 ft (50 m)	48.7	16.7
80 ft (25 m)	37.6	13.4
40 ft (12.5 m)	20.5	11.2

Source: *TCRP Report 13* (*1*).

Exhibit 8-51
Suggested AGT Separation Calculation Default Values

Default Value General Method	AGT	Term	Description
20 ft	20 ft	P_e	positioning error
650 ft	**165 ft**	L	**length of the longest train**
35 ft	**0 ft**	d_{eb}	**distance from front of train to exit block**
75%	75%	f_{br}	braking safety factor — % of maximum braking rate
2.4	**4**	b	**train detection uncertainty constant — fixed block**
1	1	b	train detection uncertainty constant — moving block
3 s	**1 s**	t_{os}	**time for overspeed governor to operate**
0.5 s	0.5 s	t_{jl}	time lost to braking jerk limitation
4.3 ft/s²	**2.0 ft/s²**	a	**service acceleration rate**
4.3 ft/s²	**3.3 ft/s²**	d	**service deceleration rate**
1.5 s	**0.5 s**	t_{br}	**brake system reaction time**
60 mi/h	**50 mi/h**	v_{max}	**maximum line velocity**
165 ft	**80 ft**	S_{mb}	**moving-block safety distance**

Source: *TCRP Report 13* (*1*).
Notes: Bold type indicates AGT default values that differ from other rail transit values. A metric version of this exhibit appears in Appendix A.

The results show that separation times with a simulated single aspect block system are two to three times longer than with the more complex—and expensive—moving-block signaling system. The moving-block results are consistent with another reference specializing in AGT train control, where typical short train AGT separation with moving-block control was cited at 15 s (*16*).

The separation range is wide and highly dependent on the train control system of the proprietary AGT system. The best method of determining the minimum train separation is from the system manufacturer or designer. Using the general methodology to determine train control separation should be a last resort when specific information is not available.

AGT headways.

The selection of a minimum headway for AGT systems should reflect the train control separation, dwell time, and any operating margin that conforms with existing operations or is suggested by the system manufacturer. The typical headway of airport systems is 120 s, with a few operating down to 90 s. Claims have been made for closer headways with some proprietary systems. Headways shorter than 90 s are possible, but may limit dwell times and constrain the operating margin. They should be considered with caution unless off-line stations are adopted. Off-line stations make closer headways possible and practical—at a price.

Loading Levels

AGT loading levels tend to be atypical of transit overall.

Loading levels of AGT cars tend to be atypical of normal transit operations. At one extreme are airport shuttles with wide cars and no or few seats where loading can reach 3 p/ft length (10 p/m length) under pressure from arriving flights. Loading diversity on airport systems fluctuates related to flight arrival times, rather than 15-min peak periods within a peak hour. After an arriving flight, three trains at 120-s headways can exceed maximum loading levels—to be followed by a number of underutilized trains.

At the other extreme are the narrow, all-seated amusement park monorails with loading as low as 0.6 to 0.9 p/ft length (2 to 3 p/m length). The peak-hour factor (PHF) on the latter type systems attains 1.0 when arrangements—and continual passenger queues—ensure that every seat on every train is occupied—in some cases, through all hours of operation.

The design capacity of non-public transit AGT systems requires consultation with the system supplier. The methodologies and calculations of this manual should only be used as a last resort—and then treated as a guideline.

Off-Line Stations

Off-line stations increase capacity.

Off-line stations maximize system capacity. They are used on several rail transit lines in Japan to achieve some of the highest throughputs for two-track rapid transit lines in the world. In North America, they are the exclusive preserve of the AGT line in Morgantown, West Virginia.

Off-line stations permit a train throughput that is partly independent of station dwell time. The throughput is that of the train control system plus an allowance for switch operation and a reduced operating margin. Morgantown and certain other AGT systems use on-vehicle switching techniques where even this allowance—typically 6 s—can be dispensed with. In theory, trains or single vehicles can operate at or close to the minimum train control separation—which can be as low as every 15 s.

Major stations with high passenger volumes may require multiple platform berths, otherwise partial dwell times must be added to the train separation times to obtain the minimum headway. The design capacity of such specialized systems should be determined through consultation with the system manufacturer or design consultant.

ROPEWAY CAPACITY

Overview

Ropeways in North America are more commonly used by private owners than by public transit agencies.

This section discusses the capacity of transit modes that are hauled by cable (wire rope). Although these modes are not widely used in North America for public transit, they are sometimes considered as modal alternatives in transit feasibility studies, and have been constructed as part of a number of private developments, particularly ski areas. In Europe, funicular railways can be found in a number of hilly urban settings, and both funiculars and aerial tramways are used for access to some remote villages inaccessible by road.

Surface modes include some of the oldest mechanized purely urban transportation systems, discounting extensions of intercity rail networks into city centers. Vehicles are either permanently attached to the rope, or can attach and detach from the rope by means of a grip mechanism. In either case, the motor driving the rope is located in a remote location, not on the vehicle itself, and the vehicle operates on a guideway. As described in Chapter 2, Mode and Service Concepts, surface ropeway modes include *cable cars, inclined planes* (funicular railways), and *cable-hauled automated people-movers*.

For the purposes of this method, two capacity categories are used: (a) reversible systems and (b) continuously circulating systems. These categories include both surface and aerial ropeway modes as members.

Reversible System Capacity

A reversible system typically provides two vehicles that are always attached to a rope and that move back and forth along the line at the same time. When one vehicle is at one terminal, the other vehicle will be at the opposite terminal. Vehicles are accelerated to line speed by increasing the speed of the haul rope and decelerated by slowing the haul rope. Passenger loading and unloading occurs while the vehicle is stopped. Modes that fall into this category are inclined planes and aerial tramways.

The line capacity of a reversible system depends on the length of the line, the line speed, and dwell times at stations. Reversible systems are usually designed with only two stations. A third station, if used, desirably should be located exactly halfway along the line so that both vehicles can be in the station at the same time. If the station is not located exactly halfway, then each vehicle will make two intermediate stops: one while at the station and one while the other vehicle is at the station.

Manufacturers claim line speeds of up to 33 to 46 ft/s (10 to 14 m/s) for funiculars and up to 39 ft/s (12 m/s) for aerial tramways. The average line speed will be somewhat less, due to acceleration and deceleration needs, and (for aerial tramways) any slowing of the line required as the carrier passes over towers.

Equation 8-16 provides the directional line capacity of a reversible system (*5, 17*).

$$T = \frac{1{,}800 N_v}{(N_s t_d) + \frac{L_l}{v_l}}$$

Equation 8-16

where

T = directional line capacity (trains/h, carriers/h);

1,800 = number of seconds in an hour, divided by two;

N_v = number of vehicles (1 or 2);

N_s = number of stops per direction:
1—two-station system,
2—three-station system, with middle station exactly halfway, and
3—three-station system, with offset middle station;

t_d = average dwell time (s);

L_l = line length (ft, m); and

v_l = average line speed (ft/s, m/s).

Continuously Circulating System Capacity

A continuously circulating system provides multiple carriers, cars, or trains that move around a route that forms a loop. Vehicles can be attached to the rope at all times (*fixed grip*) or can be attached and detached as needed (*detachable grip*).

See Chapter 2 for descriptions and illustrations of the various ropeway modes.

The concept of moving at high speed along the line and detaching from the line at stops and stations is shared by all detachable-grip modes, including detachable-grip aerial lifts, funitels, cable-hauled automated people movers (APMs), and cable cars. At stops or stations, passenger loading takes place while the vehicle is stopped (cable cars and some APMs), or while moving at creep speed (0.8 ft/s or 0.25 m/s). Manufacturers claimed line speeds range up to 20 ft/s (6 m/s) for detachable-grip gondolas, to 23 ft/s (7 m/s) for funitels, and 26 ft/s (8 m/s) for cable-hauled APMs.

Fixed-grip modes do not detach from their haul rope. Fixed-grip ski lifts load and unload passengers at line speed, but for other applications, the rope must be brought to either a full stop or creep speed at stations. To minimize the number of stops that passengers must make between stations, many fixed-grip gondola systems are designed as *pulse* systems, with three or four carriers attached in a series. At the station, all of the carriers in the series can be loaded and unloaded at the same time, thus minimizing the number of intermediate stops and improving overall travel speeds. Fixed-grip gondolas have a maximum claimed line speed of 23 ft/s (7 m/s).

The line capacity of a continuously circulating system depends on the average line speed and the spacing between carriers. For APMs, which can have multiple stations, dwell time is used to develop the minimum safe spacing between trains, following the generalized rail transit capacity method described above. Platform doors are often used both for safety (keeping passengers from falling onto the tracks or between cars of a train), and to control dwells, by keeping late-arriving passengers on the platform from holding the train doors open. For fixed-grip aerial lifts, dwell time is incorporated in the average line speed. Dwell time is not a factor for detachable-grip aerial lifts and funitels, as the carriers circulate through the station at a constant, low speed, without stopping. Equation 8-17 provides the directional line capacity of a continuously circulating system (*5, 17*).

Equation 8-17

$$T = 3{,}600 \frac{v_l}{d_c}$$

where

T = line capacity (trains/h, cars/h, carriers/h);

v_l = average line speed (ft/s, m/s); and

d_c = average carrier/train/car spacing on the line (ft/carrier, m/carrier).

Person Capacity

Manufacturers' stated capacities typically do not account for loading diversity.

Manufacturers of ropeway systems tend to state *theoretical* person capacities, based on the maximum number of people that can be carried over the course of an hour, assuming all passenger space within each vehicle is occupied. For some applications that may experience constant queues, such as ski areas, this may be a reasonable assumption. However, for public transit use, as well as any other application where minimizing passenger wait time is desired, a peak-hour factor should be applied. The

PHF accounts for the system's inability to fill every seat in every vehicle, as some capacity is reserved to handle surges in passenger demand.

The person capacity of an aerial ropeway cannot be easily increased, except for gondola systems designed with future expansion in mind.

Changing the person capacity of aerial ropeway systems is difficult, because the infrastructure (e.g., towers, rope size, vertical clearances) is designed around a particular number and size of carriers. Changing the carrier size typically requires major changes to the infrastructure. However, it is possible to design a gondola system for a larger number of carriers than will be used initially. This reduces initial capital costs and allows the capacity to be better matched to demand, as additional carriers can be added later as needed, up to the maximum number for which the system was designed.

Because the number of carriers used on detachable-grip systems can be varied by the operator, the person capacity of these systems can be adjusted over time by adding additional carriers. In this case, consideration should be given to differentiating between capacities that can currently be achieved with a given number of carriers and the maximum capacity that could be achieved.

The sizes of the cabins used by the various modes addressed in this section vary greatly. Once a particular cabin size is selected, it is difficult—if not impossible—to add person capacity by using larger carriers without rebuilding much of the system. Other infrastructure elements (e.g., towers, platforms, clearances) are designed around a particular carrier and may not be able to accommodate a larger carrier. Exhibit 8-52 provides typical ranges of cabin sizes for each mode.

Exhibit 8-52
Typical Cabin Sizes of Ropeway Modes

Mode	Capacity (p/car)	Comments
Surface Modes		
Inclined plane/funicular	20-175	Two-car trains possible
Automated people mover	30-140	Multiple-car trains possible
Aerial Modes		
Aerial tramway	20-180	Double-decked at upper limit
Gondola	4-15	
Funitel	24-30	

Source: Manufacturer data.

6. APPLICATIONS

DESIGNING FOR FUTURE GROWTH

The long-term capacity of a rail system is the design capacity achievable when the system is saturated and provided with a full complement of rolling stock. It is not the capacity that a rail transit line will provide on opening day or reach after a decade. Instead, it is the capacity after decades of growth.

Design for mature capacity.

A difficult question is for what ultimate capacity a rail transit system should be designed. Certain transportation models can predict passenger demand for several decades ahead. However predictions beyond 10 to 15 years are of decreasing accuracy—particularly in areas without an existing rail transit system or good transit usage. The resulting uncertainty makes the modal split component of the model difficult to calibrate. Even a 10- to 15-year projection period can introduce some uncertainty into the results.

It generally is preferable to base an estimate of future required capacity on assumptions about the ultimate future size of the ridership market to be served, even if these assumptions are relatively generalized. When modeling does not provide a reasonable or believable answer, or where information on future ridership potential is simply unavailable, it is possible to fall back on an old rail transit rule of thumb, namely, to design for three times the initial mature capacity. Mature capacity occurs 5 to 10 years after a system opens, when extensions and branches are complete, modal interchanges—bus feeders and park-and-ride—have matured, and some of the rail transit-initiated land use changes, including development and densification around stations, have occurred.

The line capacity determined from this manual can be used to establish the train and station platform lengths and the type of train control that will allow this long-term demand to be met—whether the demand is obtained from a long-range model or by rule-of-thumb. This long-term demand may be 30 to 50 years ahead. If this suggests that 600-ft (180-m) trains and platforms will be required, it does not mean they have to be built initially. Stations can be designed to have platforms expanded in the future. However, underground stations should have the full length cavity excavated—otherwise it can be difficult and expensive to extend platforms while the rail line is operating.

PLANNING-LEVEL ANALYSIS

The planning procedures require two main inputs: (a) train control system and (b) train length.

With the relative uniformity in the performance of electric multiple-unit trains in urban rail transit service, a simple procedure can be applied to estimate a range of peak-hour passenger capacities for grade-separated lines at their maximum load point.

The necessary choices are only two, the type of train control system and the train length. The range is provided by assigning (a) a range centered around a typical dwell time plus operating margin and (b) a small loading range centered around a comfortable peak hour average space per passenger of 5.4 ft^2 (0.5 m^2). As this is a peak-hour average, no peak-hour factor is required.

This procedure assumes system and vehicle characteristics that are close to the industry norms listed in Exhibit 8-53. It also assumes that there are no speed-restrictive curves or grades over 2% on the approach to the station with the longest dwell time, and that the power supply voltage is regulated within 15% of specifications. Finally, it assumes an adequate supply of rolling stock, and a system design that ensures that junctions (including multiple line merges) and turnbacks will not be the capacity constraint.

If any of these assumptions are not met, then the planning procedures should be used only as guidelines and the more detailed procedures in Section 5 should be used to develop a planning-level estimate of capacity.

Key assumptions are:

- *Flying junctions or no junctions,*
- *No turnback constraints,*
- *The sum of dwell time and operating margin at the critical station is no more than 70 s,*
- *No speed-restrictive curves or grades on the critical station approach,*
- *Adequate supply of rolling stock.*

Exhibit 8-53
Rail Transit Performance Assumptions for Planning Applications

Description	Default
Grade into headway critical station	< ± 2%
Distance from front of train to station exit block	<35 ft (<10 m)
% service braking rate	75%
Time for overspeed governor to operate	3.0 s
Time lost to braking jerk limitation	0.5 s
Service acceleration rate	4.3 ft/s^2 (1.3 m/s^2)
Service deceleration rate	4.3 ft/s^2 (1.3 m/s^2)
Brake system reaction time	1.5 s
Maximum line velocity	60 mi/h (100 km/h)
Dwell time	35-45 s
Operating margin	20-25 s
Line voltage as % of normal	>85%
Moving-block safety distance	165 ft (50 m)
Average peak-hour passenger loading level—light rail	1.5 p/ft length (5 p/m)
Average peak-hour passenger loading level—heavy rail	1.8 p/ft length (6 p/m)
Maximum car length—light rail	95 ft (29 m)
Maximum car length—commuter rail	85 ft (26 m)
Maximum car length—heavy rail	75 ft (23 m)

Source: *TCRP Report 13* (*1*).

Grade-Separated Rail Capacity

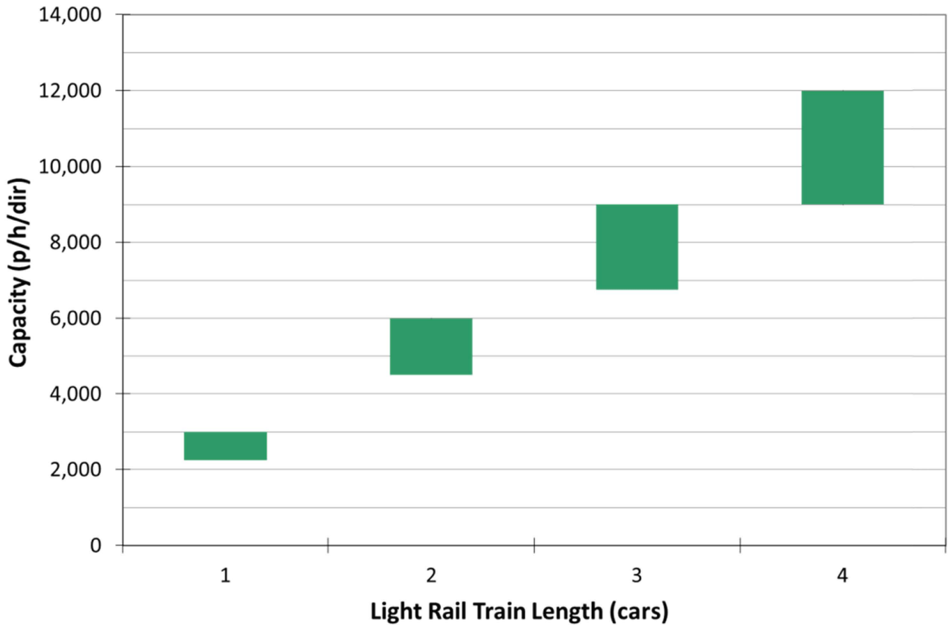
For the purposes of this procedure, grade-separated rail includes all heavy rail, portions of light rail that operate on grade-separated rights-of-way, electric commuter rail lines operating on their own trackage, and AGT systems with characteristics similar to light or heavy rail transit. The capacity of other types of AGT and commuter rail lines cannot be determined with this planning-level method and either the more detailed method presented in Section 5 or simulation (discussed later in this section) should be used.

Systems Designed for Economy

Systems that are designed economically for the minimum *planned* train headway, rather than the minimum *possible* train headway—typically, light rail systems—will design the signal and power system to accommodate this minimum planned headway.

In these cases, line capacity is directly related to the signaling constraint built into the system (assuming no significant single-track sections), and person capacity is then directly related to the line capacity and the train length. Exhibit 8-54 shows the hourly directional person capacity of light rail systems designed for a particular minimum planned headway and a particular maximum train length.

Exhibit 8-54
Capacity of Light Rail Systems Designed for Minimum Planned Headway



Note: Signal system design headway ranges from 3 min (upperbound) to 4 min (lowerbound).

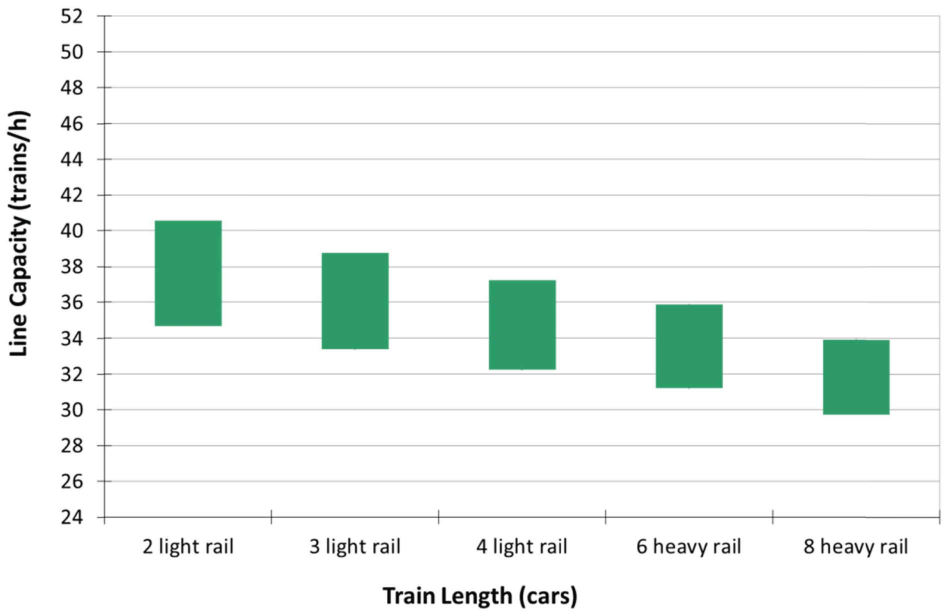
A three-aspect fixed-block system typically can support no more than 30 trains per hour—and less if a line has flat junctions or a station with extended dwell times.

Systems Designed for Maximum Throughput

As described in Section 3, three types of signaling systems are possible: fixed block, cab, and moving block. New systems that are designed for maximum throughput capacity (i.e., operations at minimum practical headway) would not use the more limited and more expensive (due to the number of signal installations required) three-aspect fixed-block signaling system. A fixed-block system may be used for systems designed for less than maximum throughput, in which case Exhibit 8-54 should be used. Consequently, the choice of train control system is limited to cab and moving-block signaling. Exhibit 8-55 through Exhibit 8-58 give line capacity and person capacity for both cab and moving-block signaling systems, based on the assumptions given in Exhibit 8-53.

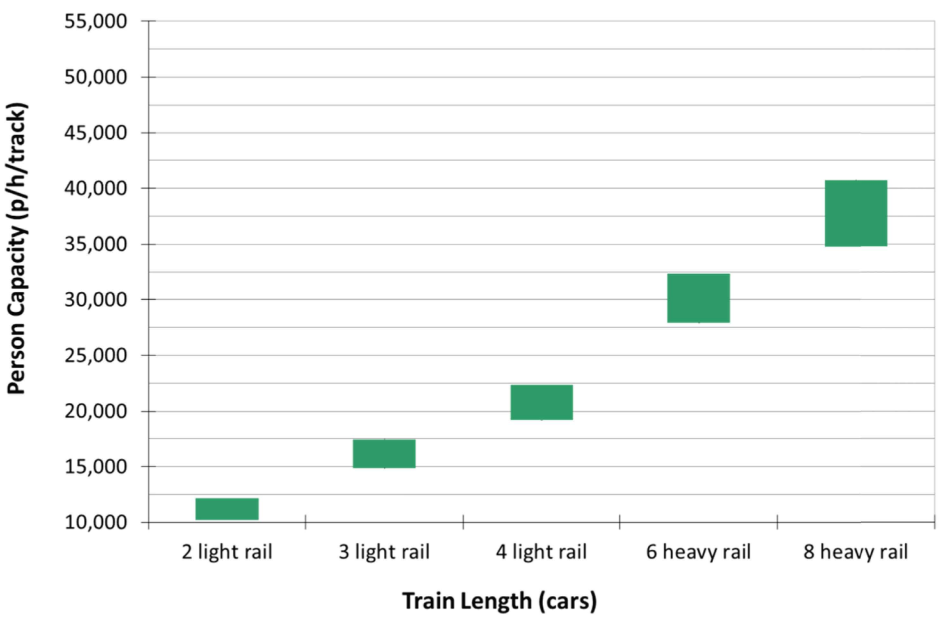
Note that with the exception of San Francisco's Muni Metro, signaled grade-separated light rail lines are rarely provided with the minimum headway capabilities represented by the capacity ranges in Exhibit 8-55 and Exhibit 8-57. Also, operating experience in North America suggests a maximum of 30 trains per hour per track for conventional rapid transit lines. It is apparent from the observed operating experience in New York and Washington that higher dwell times at critical stations prevent the achieving of capacities significantly greater than 30 trains per hour.

Exhibit 8-55
Grade-Separated Line Capacity—Cab Signaling



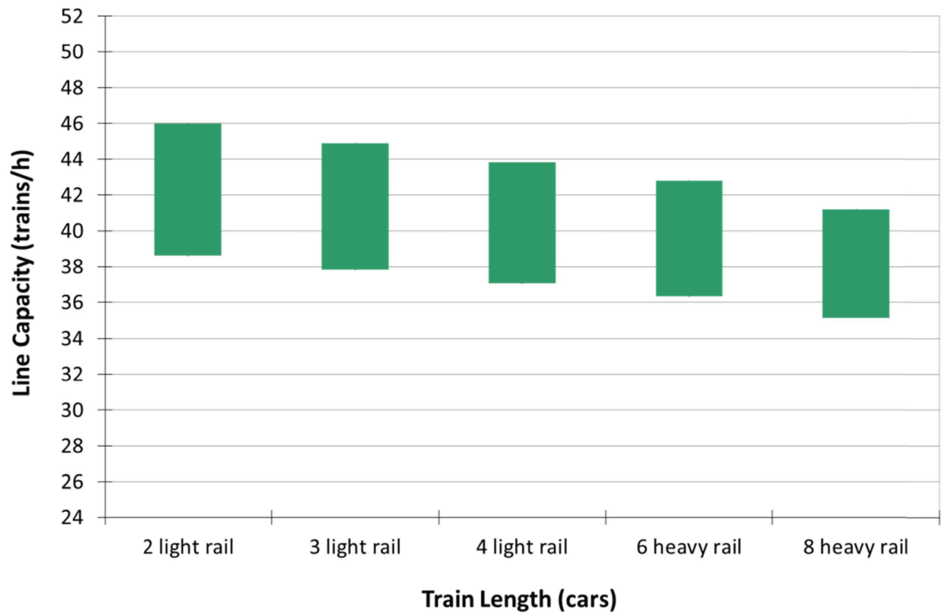
Note: Combination of dwell time and operating margin ranges from 55 s (upperbound) to 70 s (lowerbound).

Exhibit 8-56
Grade-Separated Person Capacity—Cab Signaling



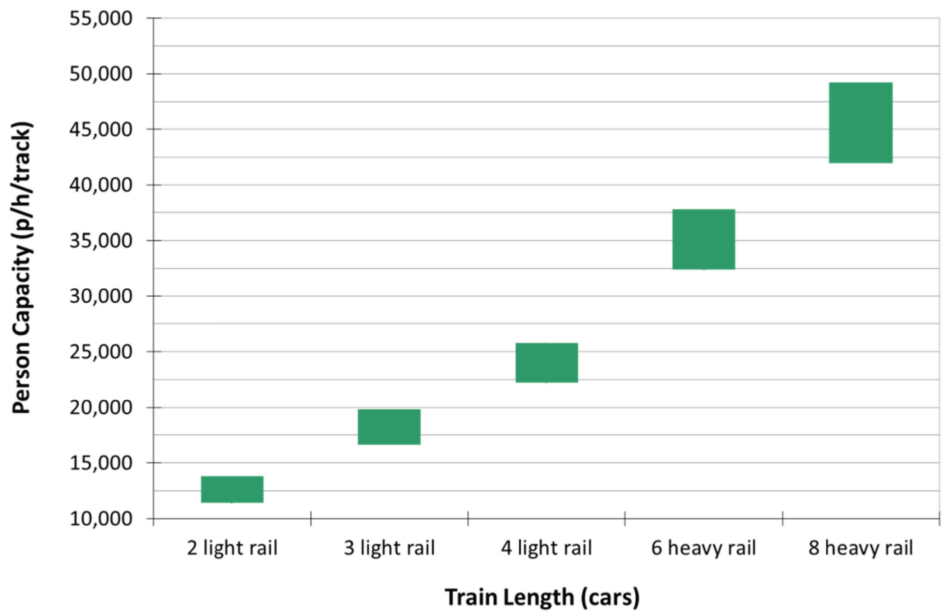
Note: Combination of dwell time and operating margin ranges from 55 s (upperbound) to 70 s (lowerbound).

Exhibit 8-57
Grade-Separated Line Capacity—Moving-Block Signaling



Note: Combination of dwell time and operating margin ranges from 55 s (upperbound) to 70 s (lowerbound).

Exhibit 8-58
Grade-Separated Person Capacity—Moving-Block Signaling



Note: Combination of dwell time and operating margin ranges from 55 s (upperbound) to 70 s (lowerbound).

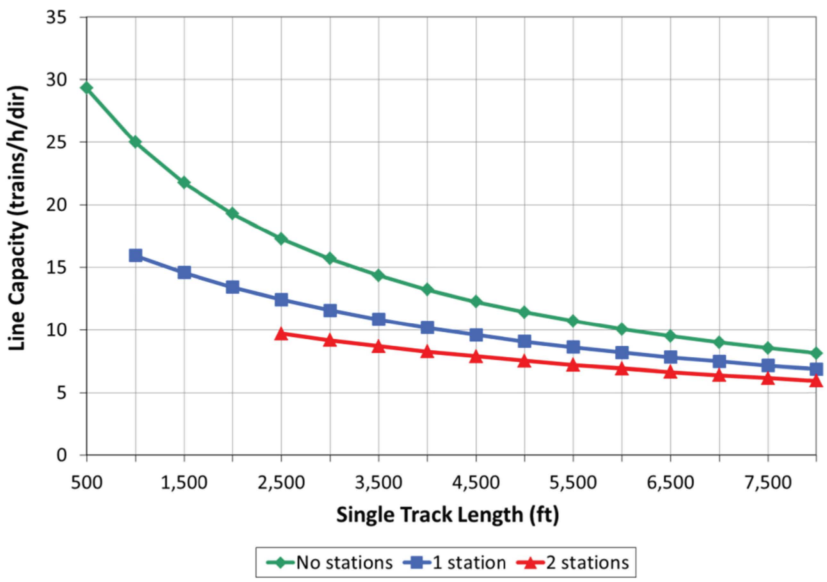
Non-grade-separated Light Rail Capacity

Light rail can operate in a variety of rights-of-way, each of which can potentially control capacity. The first of these types, grade separated, was addressed above. The remaining types—single track, exclusive lane, and private right-of-way with grade crossings—are covered in this subsection. The lowest capacity of the various right-of-way types along the line will control the overall capacity.

Single Track

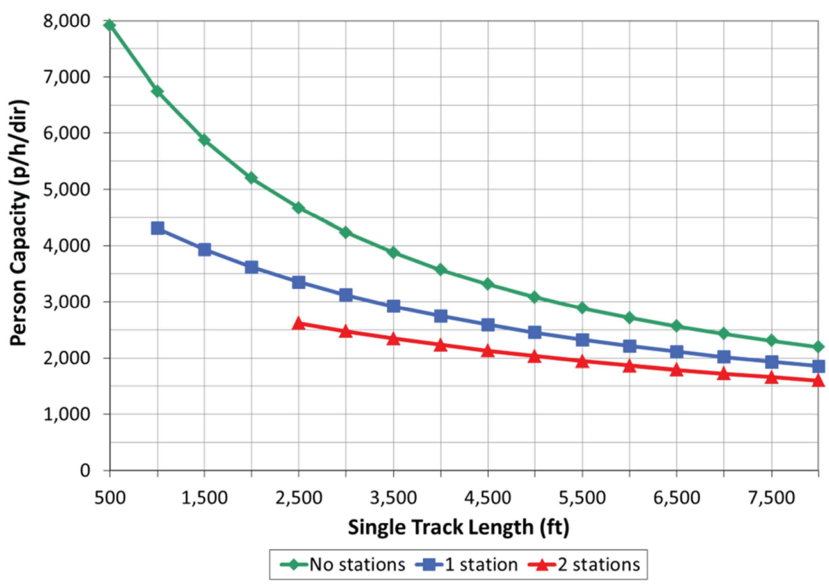
Single-track sections with two-way operation will typically be the capacity constraint when they are present. Exhibit 8-59 provides the directional line capacity of single-track sections of various lengths, with and without stations within the single-track section. Exhibit 8-60 provides the directional person capacity. The exhibits are for two-car trains. The line capacity for longer trains will be slightly lower for short single-track sections with no stations (approximately 5% lower for a 650-ft [200-m] long section), but nearly the same for long sections, or when stops are made within the single-track section.

Exhibit 8-59
Single-Track Line Capacity—
Two-Car Light Rail Trains



Notes: Assumes 35-mi/h speed limit, 180-ft train length, 20-s dwell time, and 20-s operating margin.
A metric version of this exhibit appears in Appendix A.

Exhibit 8-60
Single-Track Person Capacity—
Two-Car Light Rail Trains



Notes: Assumes 35-mi/h speed limit, 180-ft train length, 20-s dwell time, and 20-s operating margin.
A metric version of this exhibit appears in Appendix A.

Exclusive Lane Operation

The minimum sustainable headway in exclusive lane on-street operation is typically twice the longest traffic signal cycle length. When cycle lengths are long and no signal priority is provided for light rail, exclusive lane operation may constrain capacity. Exhibit 8-61 provides the line capacity for a variety of signal cycle lengths, and Exhibit 8-62 provides the corresponding person capacity. These exhibits are not applicable to streetcar operation where more than one streetcar can occupy a station at a time or where streetcars operate in mixed traffic.

Exhibit 8-61
Light Rail Line Capacity—Exclusive Lane Operation

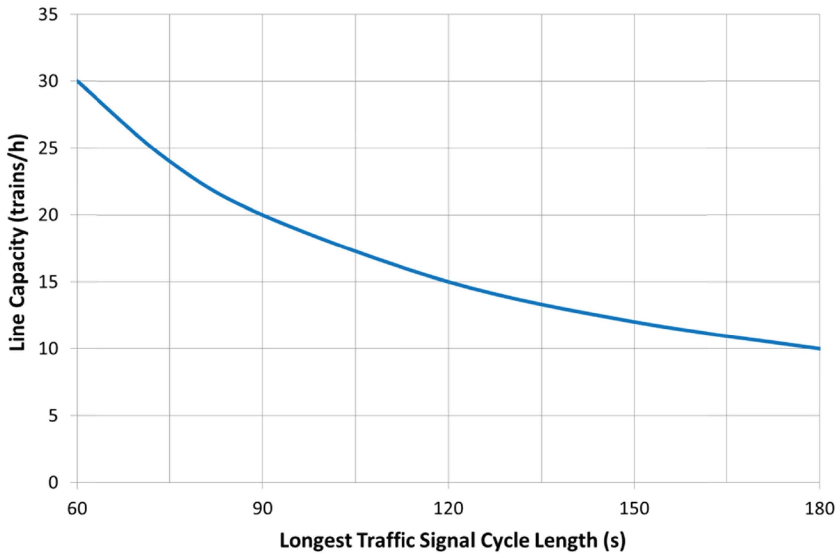
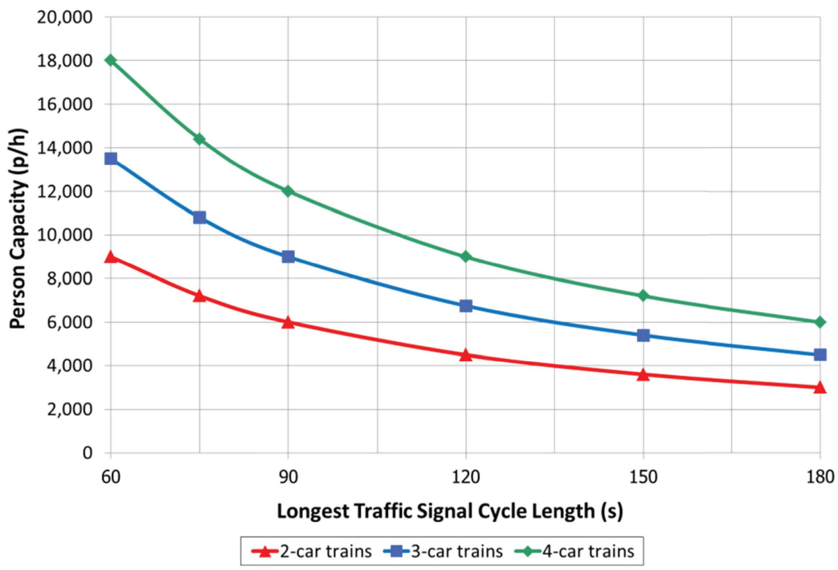


Exhibit 8-62
Light Rail Person Capacity—Exclusive Lane Operation

Exclusive lane, on-street operation is unlikely to be the capacity constraint when traffic signal cycle lengths are relatively short.



Private Right-of-Way with Grade Crossings

This category includes railroad-type operations, with street crossings controlled by gates, and operations within street medians, with street crossings controlled by traffic signals. When trains have full preemption of traffic (e.g., at gated crossings, or when full signal preemption is provided at traffic signals), the grade-separated capacity charts presented earlier may be used. Additional dwell time may be needed when station exits are located near grade crossings and preemption of the crossing is not allowed until passenger movements have ceased and the train is ready to leave the station. When trains do not have full preemption of traffic, use the exclusive lane charts above.

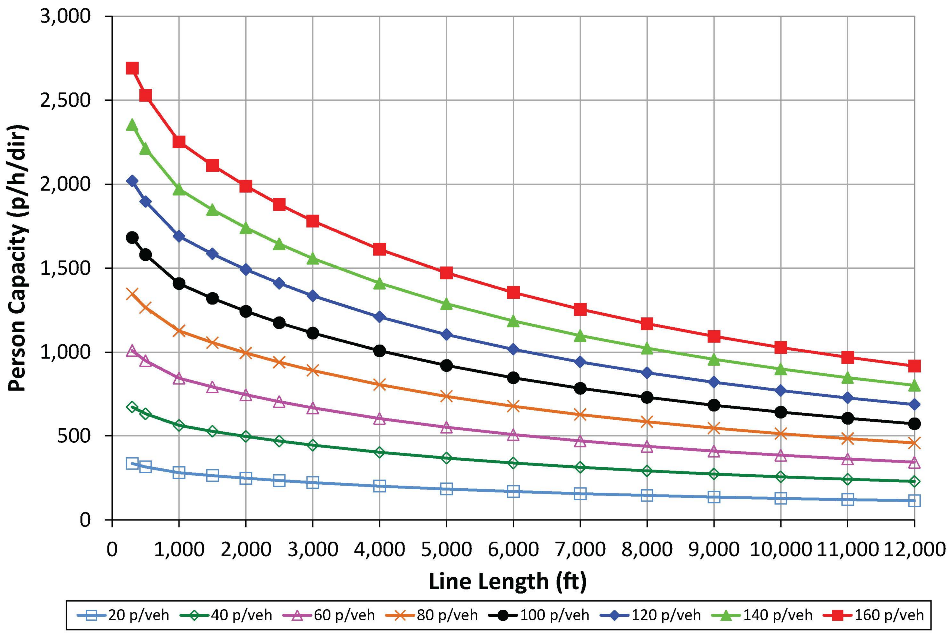
Ropeway Systems

As discussed in Section 5, ropeway systems can be classified into two categories for capacity analysis: (*a*) reversible systems, where one or two vehicles shuttle back and forth along a line, and (*b*) continuously circulating systems, where vehicles or cabins circulate around a loop. Reversible modes include aerial tramways and inclined planes. Circulating modes include gondolas and cable-hauled automated people movers.

Reversible System Capacity

The line capacity of a reversible system is dependent mainly on the length of the line and the speed at which a vehicle (train or cabin) can move from one end of the line to the other. Acceleration and deceleration delays and station dwell time are also major components of line capacity for shorter systems. Exhibit 8-63 provides the person capacity of reversible systems of various lengths and vehicle sizes, assuming two-vehicle operation and line speeds toward the upper end of modern aerial tramways and inclined planes.

Exhibit 8-63
Reversible Ropeway Person Capacity

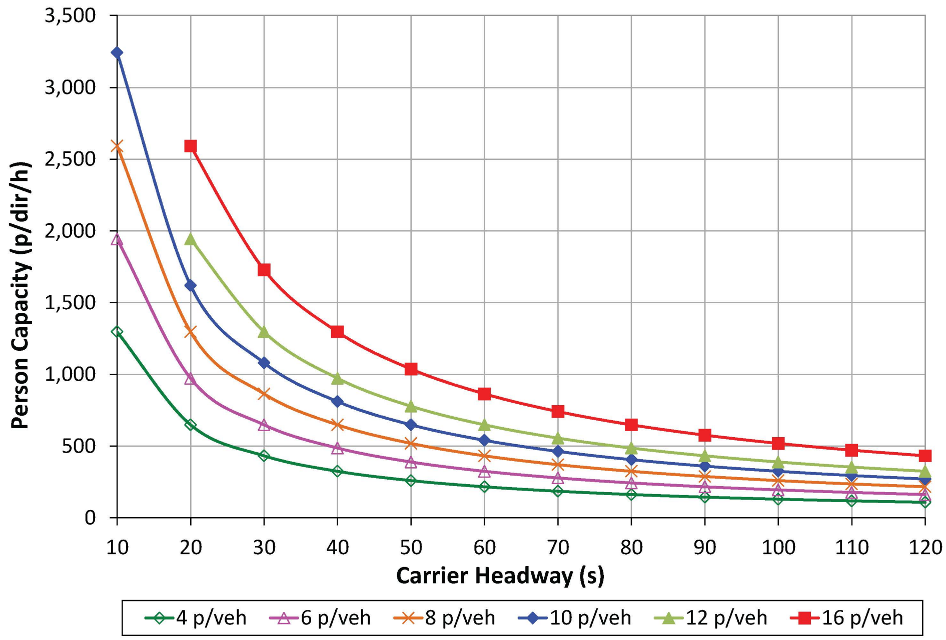


Note: Assumes 33-ft/s line speed, 0.66-ft/s^2 acceleration, two-vehicle operation, no intermediate stations, 150-s dwell time, and 0.90 PHF.
A metric version of this exhibit appears in Appendix A.

Continuously Circulating System Capacity

The line capacity of a continuously circulating system is dependent solely on the spacing of carriers or vehicles on the line. Person capacity, therefore, is simply a function of line capacity, vehicle size, and passenger arrival characteristics. Exhibit 8-64 provides the person capacity of detachable-grip gondola systems with different cabin sizes and headways.

Exhibit 8-64
Gondola Person Capacity



Note: Assumes a peak-hour factor (PHF) of 0.90.

TRANSIT OPERATIONS PLANNING

Analytic Needs for Operations Planning

Transit operations planning is performed to fulfill any of several analytic needs:

- Travel demand analysis: service headways and realistic journey times are key inputs to travel demand forecasts;
- Capacity analysis: ensuring that transportation infrastructure capacity is sufficient for the planned level of transit service;
- Analyzing and resolving current operating problems, and identifying and avoiding potential future operating problems;
- Operations and maintenance cost estimating;
- Energy consumption analysis;
- Operational input to air quality and noise/vibration impact assessments;
- Operational input to life safety analysis and transit security assessments;
- Operating crew and other staffing requirements;

- Transit vehicle fleet requirements (and associated capital costs): determining the most appropriate equipment type and the number vehicles required (including spare/out-of-service allowances); and
- Vehicle storage and maintenance facility requirements (and associated capital costs).

Service and Operating Plans

A *service plan* can be prepared for an individual transit line or for an entire network and provides a summary description of the basic characteristics of the service. It also defines the level of service offered to riders and quantifies the essential service parameters, including a description of the route, terminal points and intermediate stations/stops, and service headways by time of day and day of week. Service plans can be illustrated with simple maps and straight-line diagrams and summarized in spreadsheet matrices.

Operating plans provide a somewhat more detailed description of the service provided, including information on service and stopping patterns (including express and skip-stop service and branch line/network operations), service frequency, the type of transit fleet utilized, the length and composition of train consists (the cars, and in the case of railroad systems, locomotives making up a train set), overall fleet requirements, assigned routings over the network and track assignments at stations and terminals, equipment cycles and turns at terminal points (the linkages between an inbound train in one direction and the corresponding outbound train in the other direction), and the timing and routing of non-revenue train movements to and from storage yards and maintenance shops. Transit system operating plans also are an important input to the development of fleet maintenance strategies, including defining the appropriate locations and configurations of depots, shops, and yards.

Operating plans can be prepared and analyzed at two levels of detail, depending upon the type of analysis required and the time and resources available.

1. Sketch-plan models are used for most planning and feasibility studies. These models are spreadsheet-based, easy to develop and use, customizable for the needs of each project, and based on relevant experience observed at the system in question or elsewhere. Sketch plan models are sufficient for developing initial estimates of fleet requirements and operating/ maintenance costs and provide the headway and transit journey time inputs needed to develop ridership estimates. Developing proposed transit schedules in public timetable format can be useful for helping to describe new or enhanced transit services to project stakeholders and the public.
2. Detailed dynamic-simulation models can be developed for transit lines or networks to enable investment-grade analysis of transit system operations used to confirm needs prior to project implementation and to support engineering design and value engineering. Operating plans and simulation models can be used to undertake detailed scheduling and run cutting and to prepare detailed running time estimates for transit services, accounting for station dwell times, vehicle acceleration and deceleration profiles, and, for road-based systems, traffic and intersection delays.

ROLE OF SIMULATION

Definition and Applicability

An operations simulation model realistically depicts train movements over a transit or railroad network, including main lines, junctions, stations, and terminals. Simulations are powerful analytic tools that accurately represent:

- The physical characteristics of the infrastructure of a transit or railroad network,
- The performance characteristics of the trains operating on the network, and
- The signaling/train control system that governs operations of trains within the network.

These models are sophisticated programs designed to realistically depict rail operations in either a planning environment or an online control center. Simulation models can be developed and used at various levels of complexity. These models realistically simulate train movements over a variety of rail networks with different levels of complexity, multiple tracks and/or routes, and variable stopping patterns. They are able to resolve complex multi-train conflicts in realistic ways. Such models have proven to be fully capable of handling complex track configurations and mixes of trains on a network.

Dynamic simulation models are in common use among the owners, operators, and planners of transit and rail systems in the U.S. Users of simulation models include:

- Class I freight railroads,
- Amtrak,
- Commuter railroads, and
- Rail transit agencies.

Operations Simulation Model Types

NCHRP Report 657: Guidebook for Implementing Passenger Rail Service on Shared Passenger and Freight Corridors (*18*) provides a useful overview of the features and role of operations simulation models.

Simple Models

In its simplest form, a simulation model is a computer program that performs a stepwise calculation of the movement of a train over a rail corridor. Using information on speed limits, grades, train acceleration and braking rates, station stop dwell times, etc., the model calculates the speed and distance traveled by the train for each time step (e.g., every 10 s). After the model has stepped along the whole corridor, it produces a tabulation of time and distance traveled, often presented graphically as a time vs. distance string-line chart.

A model that performs this calculation for a single train moving over a rail corridor is usually known as a train performance calculator (TPC), because it calculates travel time without interference from other trains operating on the corridor at the same time. TPCs often have additional features, such as an ability to calculate energy used or fuel consumption. Single-train train performance calculations are used to determine required rail corridor upgrades and to estimate travel times before the interference effects from other trains and other typical operating delays are taken into account. For initial planning, it is customary to pad the minimum trip time by around 10 percent to estimate a practical trip time. This type of calculation can be used to investigate such questions as the reduction in journey time from increasing top speed from 79 mi/h to 110 mi/h, or from adding or omitting station stops.

Complex Models

The more complex version of a train operations simulator performs a simultaneous calculation of all train movements on the corridor, taking into account signal system characteristics, train priorities, temporary slow orders, and typical dispatcher decisions over where trains should meet or overtake each other. At their most complex, the multi-train simulations closely reproduce how a real rail corridor would be operated, taking into account all the variations in individual train performance and other operating constraints and variations. Results are usually presented as the calculated trip time for each train compared with minimum time with no interference from other trains, slow orders, etc. The difference is reported as a delay. Operation over the corridor can also be represented on a string-line chart (Exhibit 8-65) or as an on-screen animation—a sped-up version of a dispatcher's display (Exhibit 8-66). Details within these graphics are too small to be seen here but are not essential for understanding the general appearance of these types of displays.

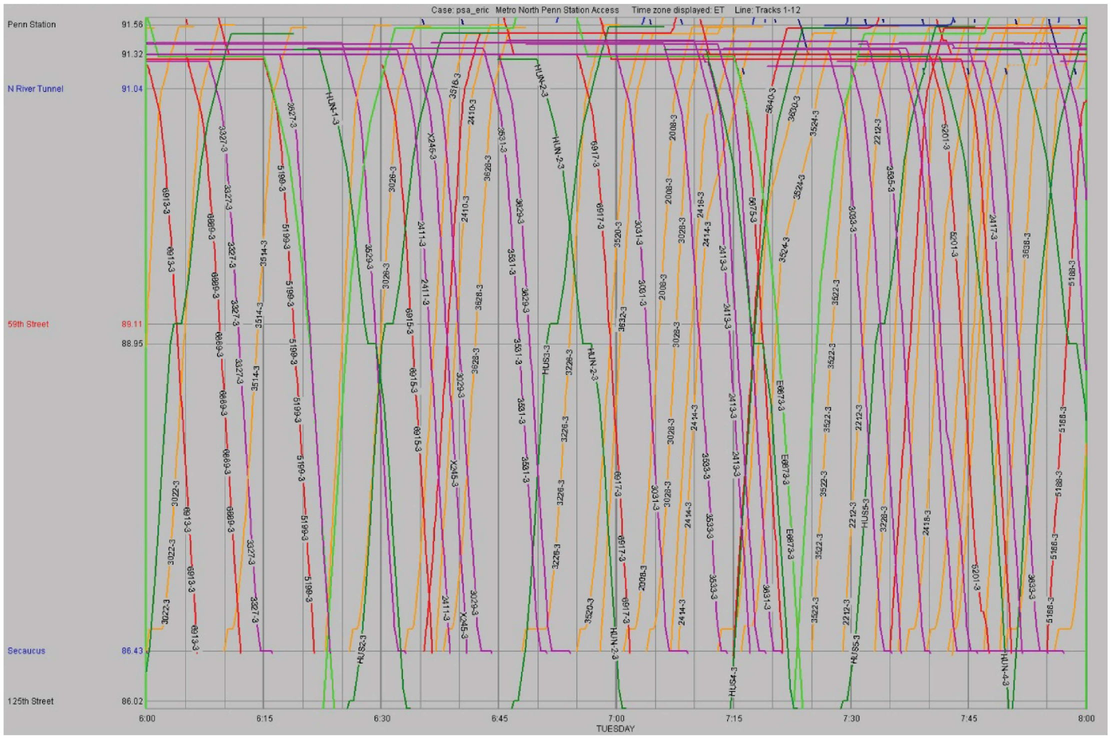
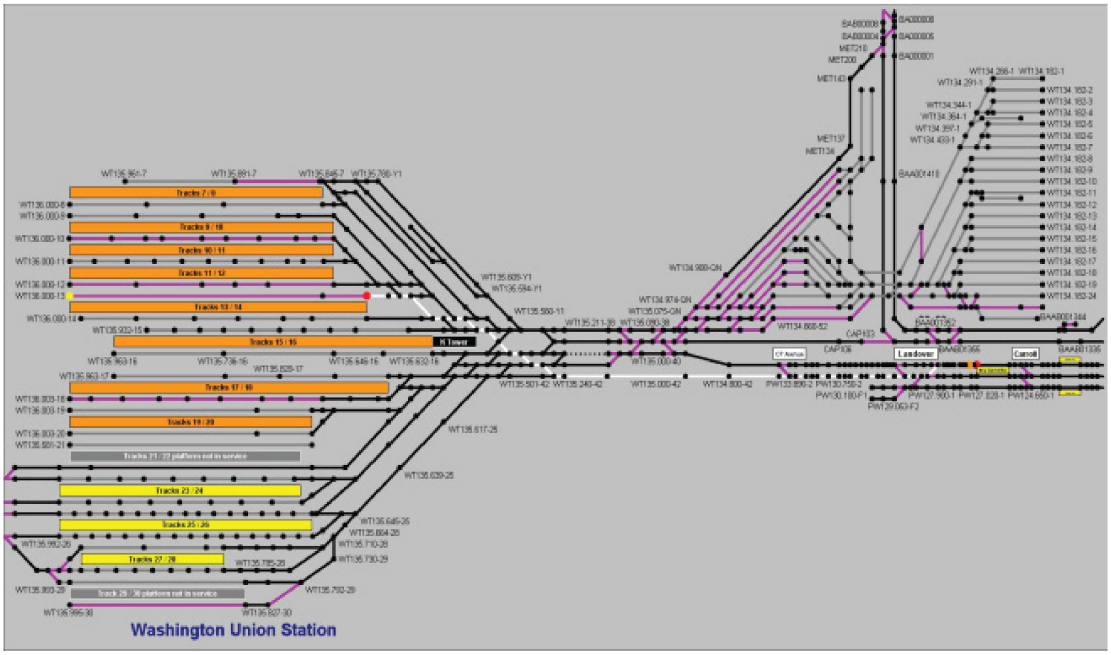


Exhibit 8-65
Example String-Line Chart: Metro North Penn Station Access

Exhibit 8-66
Example Dynamic Dispatch Animation: Washington Union Station



A single run of a corridor operations simulation will only represent operations under one set of input conditions. Railroad operations are subject to a variety of random and planned disruptions to normal operation, including planned and unplanned track maintenance, delays at stations, and delays caused by events elsewhere on the railroad. Freight train operations are not normally conducted with great precision, and even scheduled freight trains are subject to variability. In addition, many through freight trains are unscheduled "extras" that run as needed and may enter the corridor at any time. Multiple model runs are used to address these variables, with results presented as average run times and delay statistics for each train, along with string charts and animations as required.

The primary use of a multi-train simulation model is to investigate needed infrastructure upgrades to an existing rail corridor, to enable it to accommodate additional passenger train trips while still meeting specified service performance requirements (train departure times, trip time, and on-time performance) and complying with any other specified constraints. The analyst will start with improvements identified using a single-train TPC (if available) and will make multiple model runs to test alternative track configurations and other improvements. The objective is to identify a cost-effective package of improvements that will meet the service requirements of all users. Given the trial-and-error process of using simulation models, the complexity of these models, and the potentially large number of alternative corridor configurations to be investigated, an experienced modeling analyst is essential. Modeling is something of an art, and a model cannot represent everything about a route. Interpreting results requires judgment, informed by experience using the model, experience with interpreting the results, and experience observing real-life outcomes.

Simulation Software Features

There are several simulation software packages available for use in railroad and rail transit operations analysis. Simulation software packages are similar in the way that they display simulation results. While timetables and time–distance charts are useful for analysis on simple networks, they often do not show conflict resolutions at a sufficient level of detail. Simulation results can be displayed in these traditional ways, but simulations offer the benefit of providing a train dispatch animation, which, with its multitude of color modes, can bring the solutions to life. With its emphasis of graphical output, the integrity of solutions can be verified and presented without spending a large amount of time examining abstract reports.

Simulation models and other operations planning methods are intended to be objective analytic tools to help develop and test rail operating scenarios and infrastructure solutions. Simulation models have a high degree of buy-in from U.S. Class I freight railroads, Amtrak, and commuter rail operators, as a rational method to plan train schedules and resolve conflicts. They also have been adopted by most rail transit operators. The end-product of these analyses typically entails demonstrating that a recommended service and infrastructure scenario can be operated without creating an unacceptable effect on other traffic.

Two important features typically provided by simulation modeling software are train performance calculations and algorithms for train dispatching and conflict resolution. These are described below.

Train Performance Calculations

The TPC included within a simulation model is a tool used for computing minimum run times for trains operating from one specified point to another over the network without interference from other trains. Experimenting with various stopping patterns, routing configurations, dwell times, and trainset technologies helps identify the most effective scheduling/dispatching solution for a particular train type with specific physical plant characteristics.

This integrated TPC utilizes accurate trainset performance specifications in addition to length, weight, etc. The TPC applies this data in combination with tractive effort curves and dynamic braking curves or brake characteristics to replicate the dynamics of each specific train traveling over the defined physical characteristics of the network.

Dispatching Logic and Conflict Resolution Algorithms

There are two principal types of logic used by dynamic simulation models to resolve conflicts between trains operating over a rail or rail transit network. These can be categorized generally as time based and event based.

Time-based models move trains through a network in real time, advancing all trains in the network in small increments, reacting to the physical characteristics of the track and train equipment, and the indications being given by the signal and train control system. The logic of the signal system is used to resolve conflicts between trains looking to occupy a section of track or a route through an interlocking at the same time. Time-based simulations are useful for modeling most railroad and transit mainline track configurations, as well as terminals where train routings and track assignments are well defined or customized to a particular pattern of operation.

Event-based models do not simply resolve conflicts between pairs of trains, but rather look globally at multi-train conflicts and resolve them as integral elements in the dynamics of the entire network. These models can be equipped with "meet–pass" dispatching logic, which is used to develop the most effective solutions to single-track operations requiring segments with a second main track, passing tracks, or passing sidings. The proven dispatching logic that is integrated with the train control system and TPC identifies conflicts and presents solutions that effectively contribute to improving the performance, reliability, and capacity of the entire rail network. A key aspect of the dispatching logic application is alternate node logic. Alternate node logic is extremely useful for yard and terminal analysis. It is capable of defining the best route and the most appropriate platform assignment for every train based on consist properties and priority, while looking holistically at the network performance.

Event-based models have been proven to be effective in analyzing large, complex networks with high train volumes. This capability is especially important in rail networks where the density and dynamics of passenger trains require a dispatching logic that effectively addresses close-headway operations and intensive interlocking routing issues.

Simulation Model Inputs

A variety of data and assumptions associated with the railroad's physical and operating conditions are necessary to construct a computer simulation model, including:

- Conceptual design drawings for the track alignment and station configurations;
- Specifications for the limits of the simulation study area;
- Assumptions for the fixed-block signal system design;
- Operating speeds including civil speed restrictions;
- Grade and curvature alignment characteristics;
- Characteristics for special trackwork geometry;
- Assumptions on operating rules, special instructions, and policies governing train movements;
- Assumptions on service specifications, timetables, schedules, headways, and station dwell times, including distinctions between peak and off-peak operations for systems heavily used by commuters;
- Trainset performance specifications for the assumed consist configuration, consist composition (number of cars), tractive effort curves, braking effort curves, and all dimension and weight characteristics;
- Assumptions on headways and number of trains per hour to be examined in the simulation;
- Definition of the time period to be analyzed, including warm-up and cool-down periods (e.g., simulating a 40-h weekday period to provide a realistic and valid sample of train movement activity over a continuous 24-h period with the network operating in an equilibrium condition; typically, the weekday 40-h sample will include a morning and an evening peak, each of which can be broken

out into separate reports and track occupancy charts for review and analysis); and

- Detailed train schedule for the time period to be analyzed.

Relevant data representing the dynamics of the proposed operation will need to be collected from key staff in the engineering, mechanical, stations, operations, and service planning departments of the transit operator or railroad. These interfaces may include line operations and control center dispatching staff.

Key characteristics of stations that are considered in simulation models include:

- Location of stations on the network;
- Station platform configuration (identification of tracks with platforms, and location and length of platforms);
- Dwell time for trains for each station; and
- Passenger processing (horizontal and vertical circulation).

Simulation Model Outputs

The modeling software provides numerous outputs that facilitate effective evaluation of performance of different train types, i.e., high speed rail, intercity passenger rail, commuter rail, and freight rail; development of optimal operating patterns and train routing; and assessment of a rail network with different levels of service. The software can assemble operating statistics at a specific "train-by-train" level, for groups of trains, or at a systemwide level. These outputs can be categorized into three major groups: statistical data, static visual representations, and dynamic visual representations.

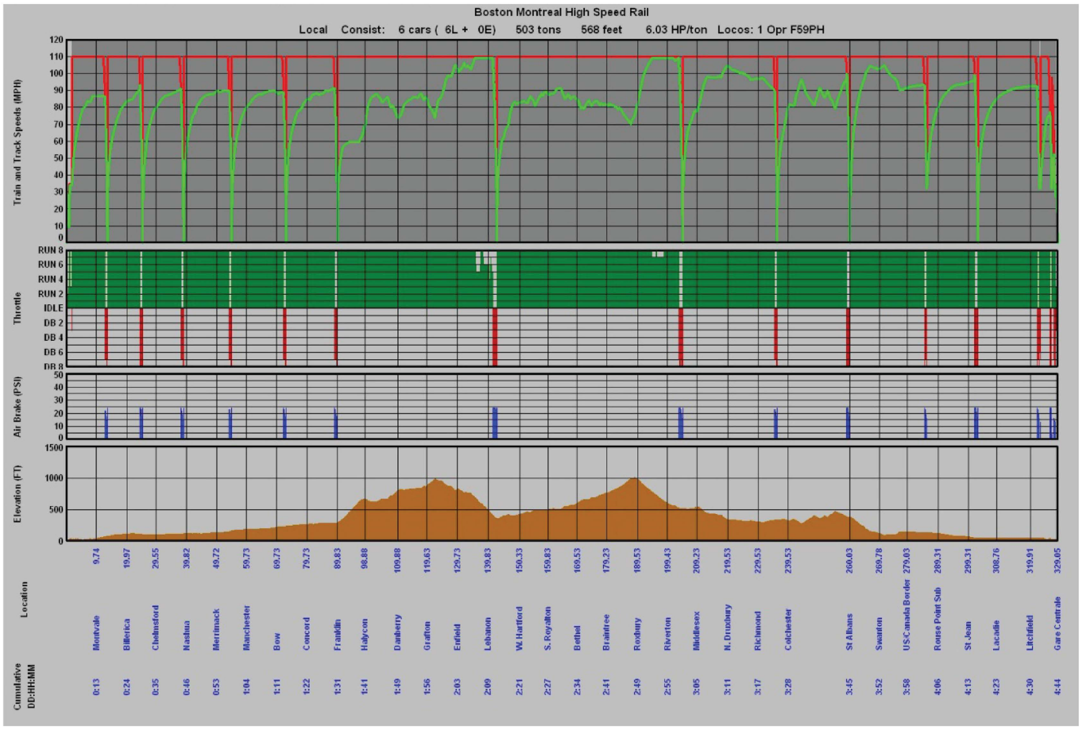
Statistical Outputs

The TPC is used to calculate minimum point-to-point travel times, including station-to-station and end-of-network to end-of-network travel times, based on the performance of a single train traveling along the route without interference from other trains. It also produces throttle and braking positions and the speed of the train at any point along the route. Analysis of the TPC speed profile(s) contributes to straightforward identification of the necessary infrastructure improvement locations when speed-increase opportunities are desired, for example, slow curves. Other types of statistical data include (but are not limited to): ideal and simulated travel times that are used to calculate delays at different levels, detailed train route descriptions, and train and car mileage data. These data typically are represented in data tables and spreadsheets.

Static Visual Representation Outputs

Speed profiles depict maximum speeds and optimal trip times from the train performance calculator (TPC). An example is shown in Exhibit 8-67. As with other screen captures in this section, the exhibit details are not legible as this scale but are not necessary for understanding the general appearance of the visual outputs.

Exhibit 8-67
Example Speed Profile



Time–distance or *string-line* diagrams depict the paths of individual trains over time and space. An example was previously given in Exhibit 8-65. The diagrams use specific color coding for each track to graphically present simulated results for operating speeds, running times, dwell times, practical minimum headways, and line capacity. The model displays the string-line chart(s) after all of the trains have been dispatched in the simulation and any operating conflicts have been identified and resolved by the model. This important analytical tool presents a static snapshot of all the trains and their interactions in the network model. A multitude of color modes makes it easy to observe trains which are stopped for delays and/or slowed due to speed reductions caused by crossing or merging conflicts. The string line graphs also provide the ability to analyze headways and identify "choke" points and capacity limitations.

Track occupancy charts graphically display the trains that occupy specific tracks in stations and yards at specific times throughout the simulation, which is very useful for identifying "slots" at station platforms, evaluating the level of utilization of platform and storage yard tracks over time, and developing equipment cycles by following the established train linkages. In addition to its more typical application for analyzing operations in yards and terminals, this chart can also be used to display the occupancy of any track segment where track utilization and train headway is the focal point of a study's analysis.

In the following track occupancy graphic (Exhibit 8-68), the times displayed for a train are from head-end arrival to rear-end departure. For example, on Track 6, eastbound Train 3516E arrives at 6:40 p.m. and departs as westbound train 3031W at 6:50 p.m. The next train on Track 6, 3520E, arrives eastbound at 7:02 p.m., retains its identity, and departs eastbound at 7:14 p.m. The trains are displayed in their proper time slots, thus providing the ability to observe train movements in either a "turn" (3516E to 3031W) or a "run-through" (3520E to 3520E) dynamic.

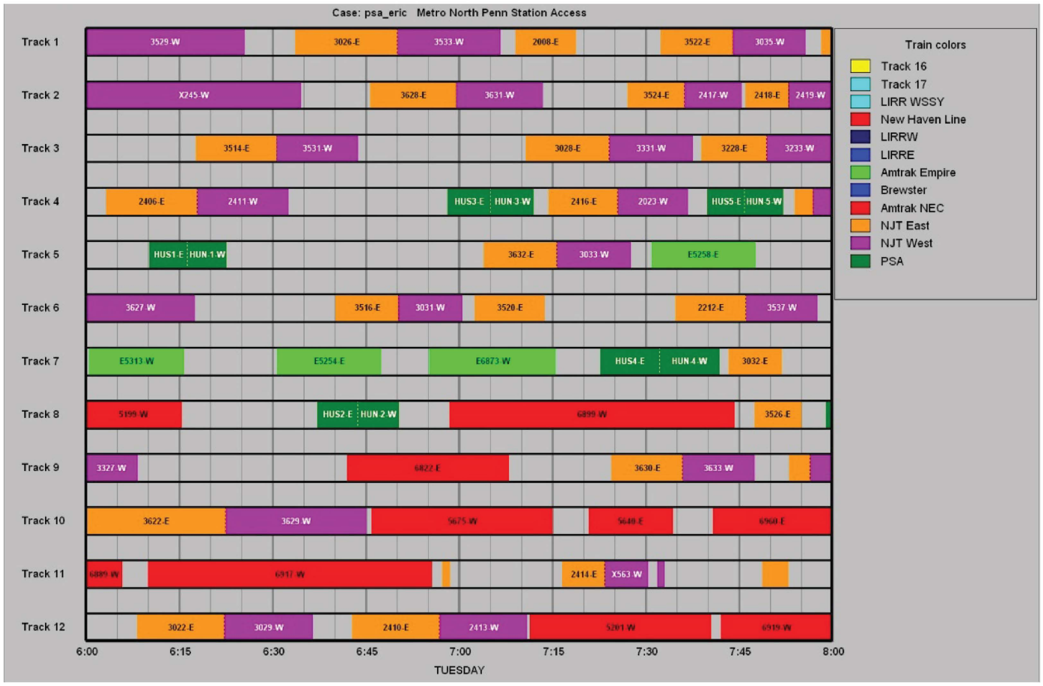
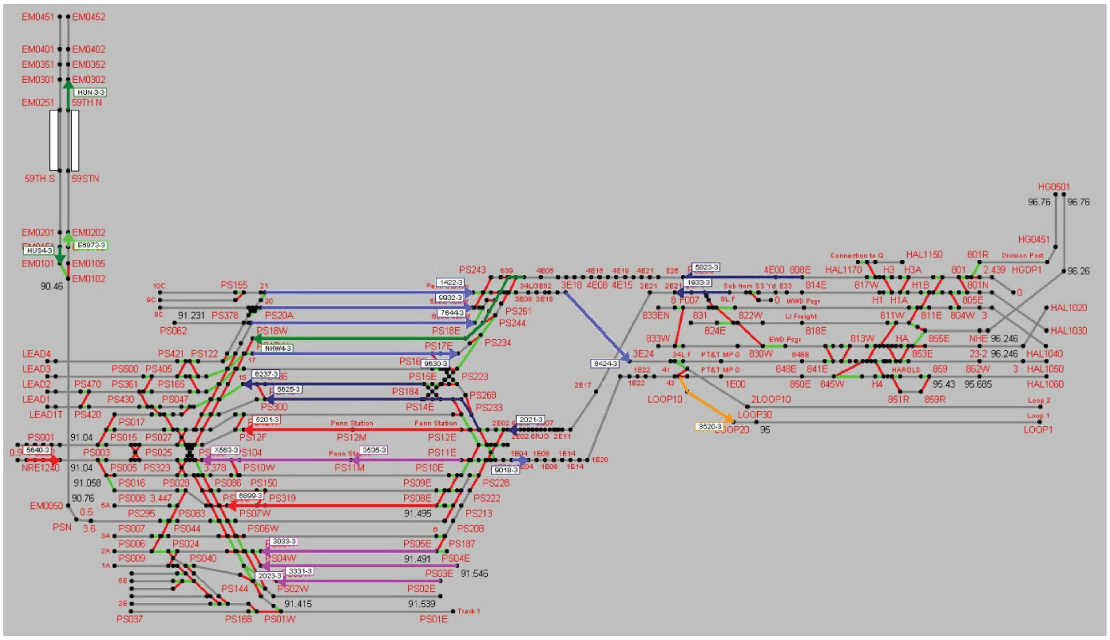


Exhibit 8-68
Example Track Occupancy Chart

Dynamic Visual Representation Outputs

Dynamic dispatch animation presents a "real time" dynamic visual representation of the railroad's operation of the railroad depicted in the simulation. It frequently is the ultimate tool for testing and validating operating plans, clearly showing the entire spectrum of train delays and meet-and-pass conflicts and how they were resolved in the network's train operations. The ability to apply the model's color mode feature provides the ability to observe the performance of an individual train or a group of trains, based on user-selected criteria. In addition, various animation display speeds can also be selected (by the user) and the simulation can be run in reverse or completely stopped at any time. The operational status of trains during animation can be displayed by the user as shown in Exhibit 8-69, where train locations and directions are shown as colored arrows. Output can be a "movie clip" file that can be reviewed by the operations analyst during the analysis process or embedded within and shown as part of a presentation to project sponsors and stakeholders.

Exhibit 8-69
Dynamic Dispatch Animation Example: Penn Station New York



APPLICATION OF SIMULATION

Base Case Model Calibration and Validation

The simulation model will first need to be calibrated and validated to reflect a realistic representation of the scheduled running times for the operations and services being analyzed. The first step in the process is to build a base case computer simulation model that includes all the parameters required by the model for the analysis, including the rail infrastructure characteristics, rolling stock characteristics, and operational service levels for the territory being analyzed. The infrastructure data will include vertical and horizontal attributes, grades and elevations, degree of curvature, speed restrictions, and maximum authorized speeds. The track layout will include all the turnouts and crossovers. Rolling stock parameters and train set performance characteristics will be defined with car dimensions and weights, locomotive tractive effort, and dynamic braking forces. A base case operating plan will be developed to reflect all the train movements to and from the station, including arrival and departure times and detailed equipment data.

The base case simulation model then will be calibrated and validated against existing conditions to verify the accuracy and realistic application of the dataset that will be used as the benchmark for further analyses. The validation phase is not considered complete until the operators of trains on the network being analyzed have reviewed and approved the base case model.

Investment-Grade Analysis

Simulation models are frequently used to support the business case for private investment in rail system infrastructure and are increasingly required to justify public investment in such systems. Simulation case studies are used to analyze the performance, operating reliability and capacity of alternative rail infrastructure solutions to respond to the assumed or desired frequency or level of service. The model is used to:

- Establish operational feasibility,
- Prove the reliability of the operation, and
- Identify order of magnitude of capital costs for infrastructure and fleet.

Developing Realistic Operating Plans—Balancing Infrastructure Investment with Operational Capacity

The simulation modeling tool eliminates the traditional practice of developing schedules and train movement alternatives based on average run times, an oversimplification that can lead to unachievable operating plans.

Arrival and departure times (as well as other parameters) are modified using the model to improve schedules and craft the most fluid train dispatching scenarios. Furthermore, as traffic density increases, the potential for conflicting train movements increases as well, resulting in exposure to delays. This is precisely where this simulation tool offers unprecedented, effective functionality.

The model simulates train movements resulting in analytical data that identifies the most effective, overall system solution. When an “excessive” number of trains are specified to operate on the network, causing congestion, the model will slow and/or delay trains as needed (either at terminals or enroute) until clear routes become available. This characteristic provides the ability to vary departure times, dwell times, and influence the dynamics of train “turns” to test schedule robustness, effectiveness of train dispatching, and physical plant capacity.

In summary, the model replicates and predicts actual train movements, accurately identifying train dispatching and routing conflicts. Each simulation case analysis delivers precise comparisons of capacity and train delay at specific (and varied) levels of train service within a specified definition of infrastructure and physical characteristics.

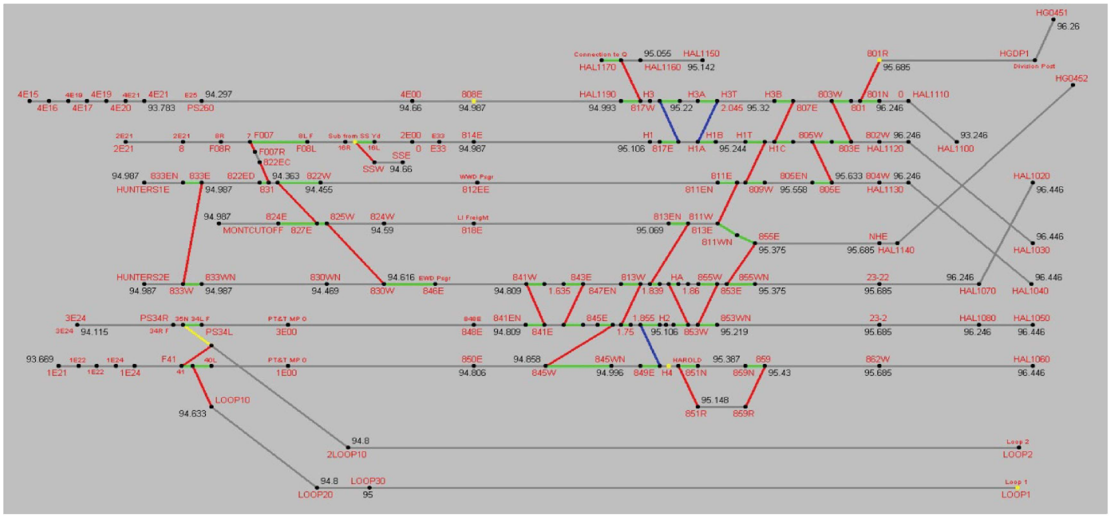
Adding Service

The effects of adding trains to a congested line or network are comprehensively evaluated using the model. The simulation tool measures the delay and performance resulting from service additions both by individual train as well as at the more aggregated levels of train type (i.e., peak vs. off-peak) and overall system network.

Interlockings and Junctions

The model is utilized to evaluate the benefits (and costs) of adding, modifying, or eliminating interlockings, either in mainline road territory or within a complicated terminal or station area, as in Exhibit 8-70. The model simulates the dynamic operating conditions and delays associated with separate or (route) segmented interlockings. The model also can be used to analyze the relative performance of at-grade versus grade-separated junctions and track connections.

Exhibit 8-70
Example Simulated Interlockings



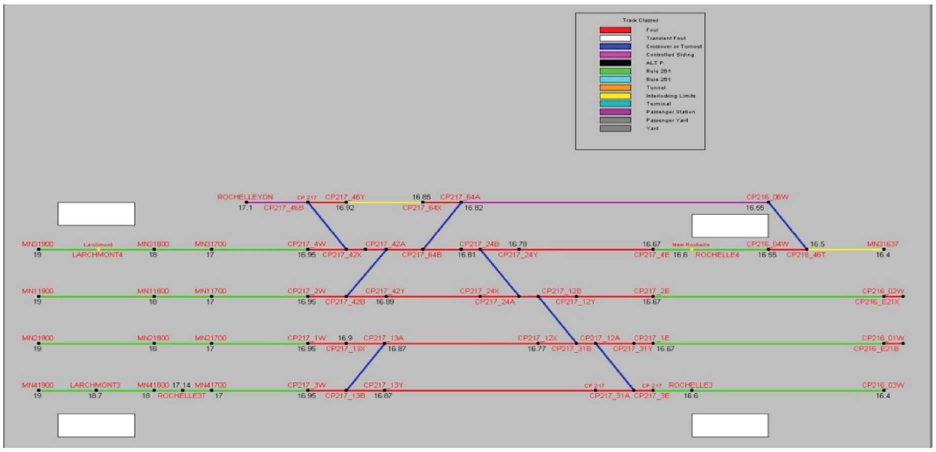
Construction Staging and Maintenance-Of-Way Windows

The model provides the ability to develop realistic construction staging plans and to schedule the most effective maintenance-of-way (MOW) time slots on busy main tracks and terminals. It graphically displays the effects of track impedances and speed restrictions on train movements. The dynamic simulation feature offers the capability to experiment with various staging scenarios and/or MOW windows to determine the best train schedules, physical plant configurations, and timeframes to construct capital improvements or perform maintenance activities.

Establishing or Moving Crossover Locations

The placement of crossovers can have dramatic effects on capacity utilization and train performance in multiple-track territories. The model provides the ability to move crossovers around the rail network and test different assumptions on speeds for diverging train movements, a significant advantage in identifying the locations that are best suited for a given set of train types and schedules. Modeling a variety of crossover configurations also contributes to developing the most effective solutions to congestion issues observed in the simulation. Exhibit 8-71 provides an example of a network graphic with crossovers.

Exhibit 8-71
Example Network Graphic with Crossovers



Single-Track Networks

Significant capital and maintenance cost savings sometimes can be realized in light density railroad and transit operations by operating service over a single main track on all or a portion of a line. These savings, however, come at the expense of capacity and operational flexibility. Simulation modeling can be used to determine the capacity of single-track segments of a railroad or transit line, identify the optimal level and type of service that can be operated over the single track segment, identify the location and length of required intermediate passing sidings, and analyze the benefits and costs of double tracking a single-track line. Examples of single-track systems where simulation modeling proved useful in defining the appropriate track configurations included the Charlotte North Corridor commuter rail line, which is envisioned as a single-track line with intermediate passing sidings, several branch lines of the Long Island Rail Road, which are double tracked over most of their length but have single-track segments near the ends of the lines, and the Baltimore Central Light Rail Line, which was converted from a single-track to double-track system.

Adding, Extending, or Removing Passing Sidings

The utility of passing sidings or long segments of main track in multiple track territory is determined by their size and location. The "ideal" location for a passing siding or additional main track segment for 30-mi/h track can be quite different than for 50-mi/h, 60-mi/h, or 80-mi/h track. The simulation model enables the user to determine whether siding or additional main track segments are of appropriate length and location for the size and speed of the trains being operated, or to identify the best train sizes and operating speeds to match a specific track configuration. Exhibit 8-72 provides an example of a network graphic for a line with passing sidings.

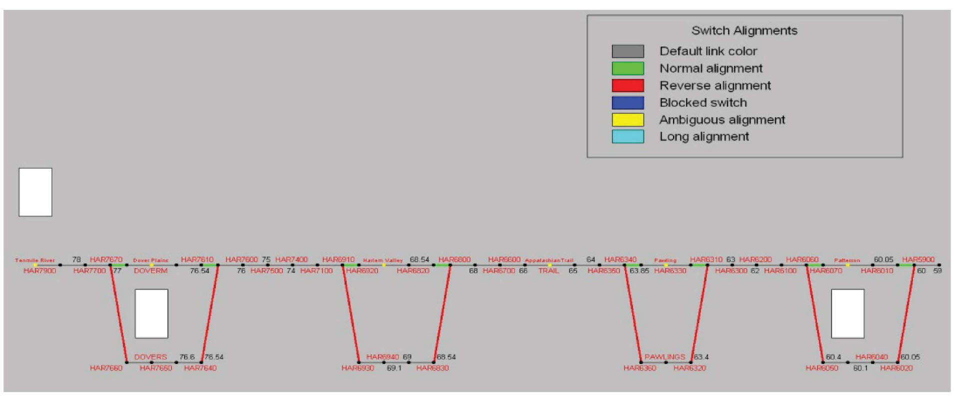


Exhibit 8-72
Example Network Graphic with Passing Sidings

SKETCH-PLANNING TOOLS

Definition and Applicability

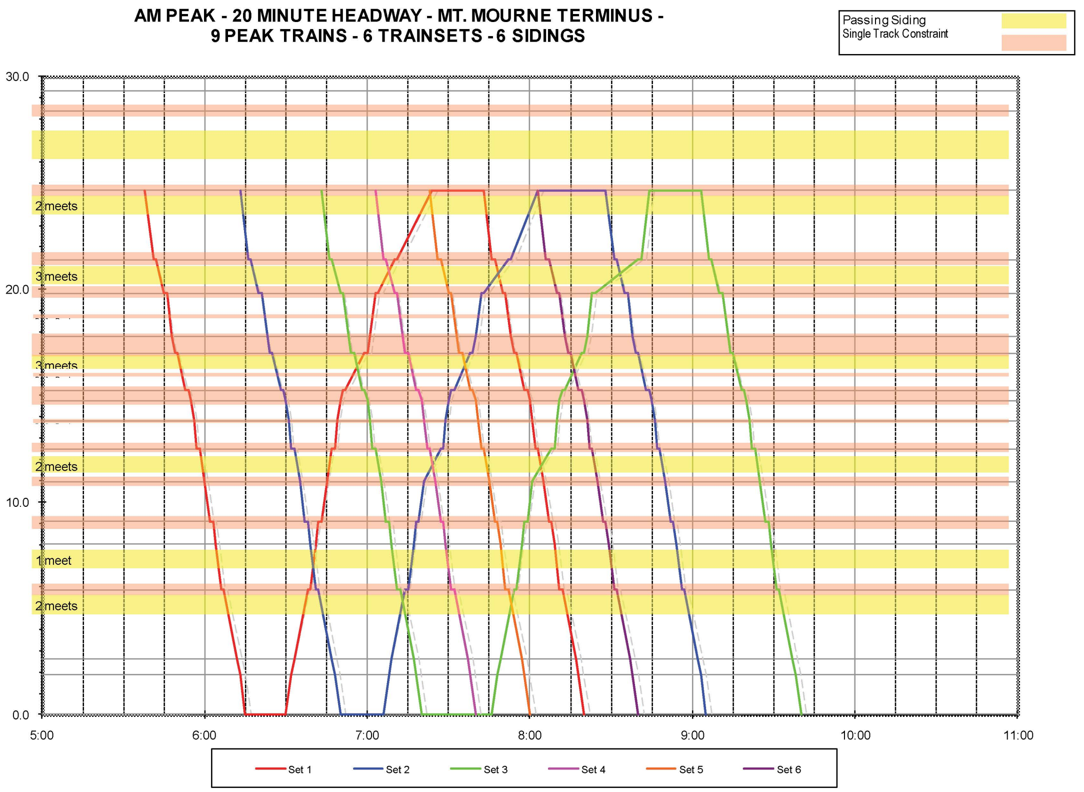
Spreadsheet-based or manual methods can be used to generate work products similar to those produced by detailed simulation models. In place of detailed train performance data based on the tractive effort of locomotives or multiple-unit transit or rail cars, simplified train movement assumptions can be captured based on average operating speeds and average train acceleration and deceleration rates. Conflict resolution on mainline track segments and at terminals can be performed manually by inspecting string-line diagrams and track-occupancy diagrams, and by adjusting train schedules, track assignments and train performance to resolve conflicts.

Sketch plan methods produce many of the same outputs as full simulations, including:

- Spreadsheet TPC,
- String-line diagrams,
- Train schedules,
- Equipment cycles,
- Track-occupancy diagrams, and
- Train loading diagrams.

Exhibit 8-73 through Exhibit 8-76 show examples of the types of outputs possible from sketch-planning tools. As before, the purpose of these exhibits is to illustrate ways that useful information for rail planning can be presented; the exhibit details (which are generally too small to be legible at this scale) are not needed for understanding the basic presentation concept.

Exhibit 8-73
Example Sketch-Plan String-Line Diagram (Charlotte North Corridor Commuter Rail)



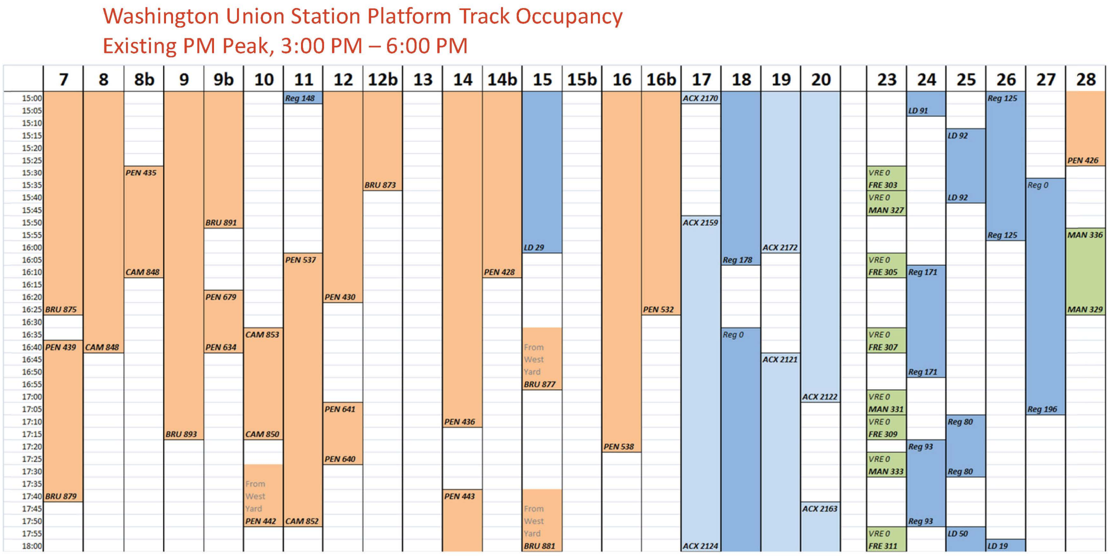


Exhibit 8-74
Example Sketch-Plan Track-Occupancy Diagram (Washington Union Station)

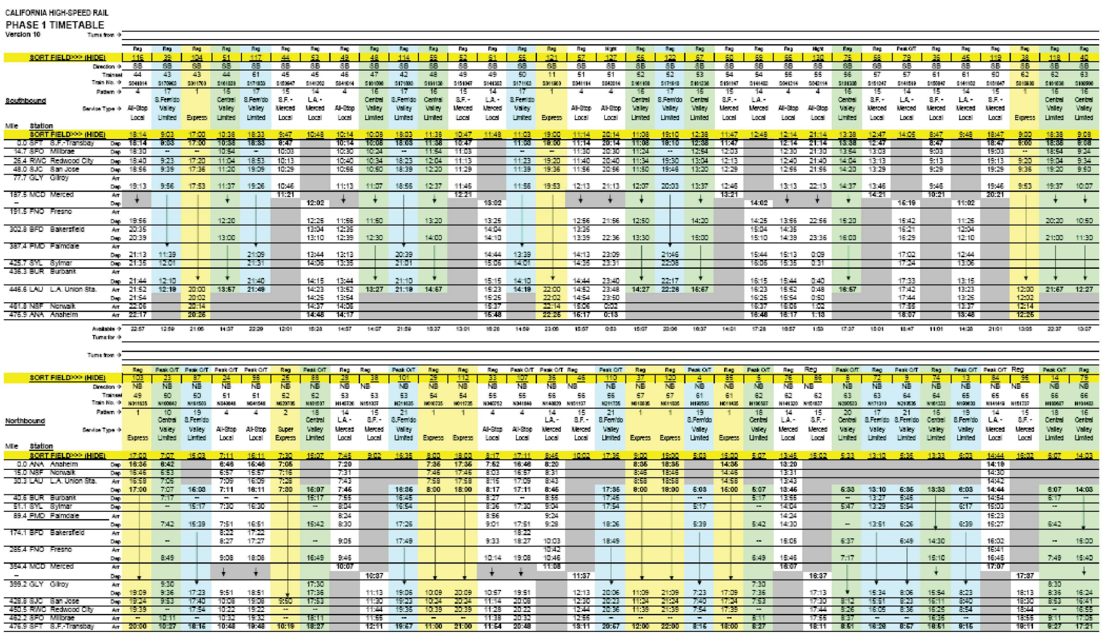
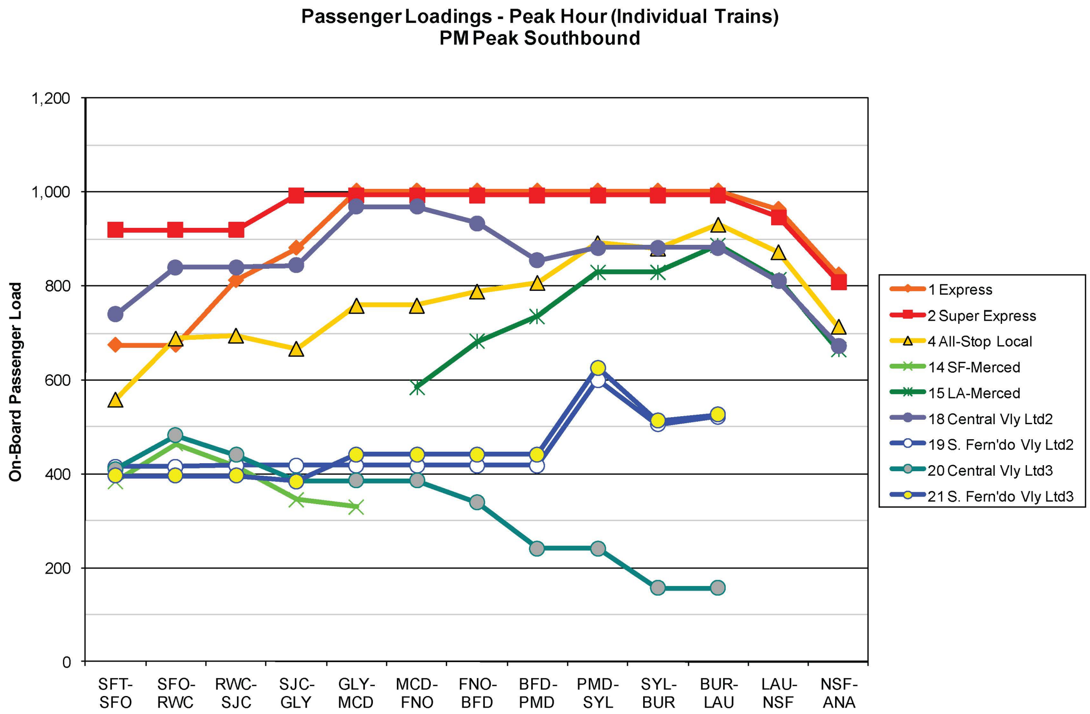


Exhibit 8-75
Example Sketch-Plan Timetable (California High Speed Rail)

Exhibit 8-76
Example Sketch-Plan Train Loading Chart (California High Speed Rail)



BEST PRACTICES FOR THE USE OF SIMULATION MODELS AND SKETCH-PLANNING TOOLS

Sketch-planning tools (including applications of TCQSM methods) are potentially sufficient for the following types of applications:

- Initial planning-level analysis of multiple modes (e.g., BRT vs. rail) or technologies within a corridor or region;
- Single networks or systems (e.g., no junctions); and
- Projects with limited resources and/or rapid turnaround times.

Dynamic-simulation models may be required for:

- Analysis of complex networks, with multiple branch lines and junctions;
- Analysis of major or complex terminals;
- Evaluation of rail or transit services that employ multiple types of stopping patterns (e.g., express, limited stop, and local service);
- Projects with schedule time and resources to support simulation model development; and
- Projects for which investment-grade analysis is required.

It also is possible to combine the use of detailed simulations and sketch-planning tools to productively analyze a wide array of alternative physical rail infrastructure configurations and/or alternative operating plans—without having to develop and run a full-scale simulation for each combination and permutation. The blended approach employs a combination of detailed simulation modeling and spreadsheet-based sketch planning to productively generate corridor and networkwide service plans and detailed operating plans for a range of alternatives. In this approach, operations planning tools are developed at two levels of detail: system-level sketch planning and detailed simulations of actual infrastructure and operating conditions. These two sets of tools are able to communicate with each other and are deployed in parallel through the various phases of analysis and alternatives screening. This approach gives the analyst the ability to study a wide array of options relatively expeditiously using the sketch-plan tools while developing a high degree of confidence in the precision and accuracy with which selected representative operating plans can be defined using the simulation tools.

Detailed simulation modeling provides the best framework for performing train performance calculations, which are the basis for estimating run times over the rail network for alignment, station and stopping pattern alternatives. The detailed simulation environment also is the best way to understand the stochastic, dynamic effects of real-world operating conditions on system performance, reliability, and practical capacity. TPC data for each potential stopping pattern are then imported into the sketch-plan model and used to generate a hypothetical timetable. Overtaking conflicts are then resolved by reviewing and manipulating string-line diagrams. Once a conflict-free schedule has been developed, additional sketch-plan modules estimate equipment turns, fleet requirements, crew schedules, operational parameters such as train and car miles, and midday and overnight storage requirements.

Standard service plan outputs (for each discrete scenario and variation, generated primarily from the spreadsheet-based sketch-plan models) include:

- Train timetables (usually for a typical weekday);
- String-line diagrams (spreadsheet generated);
- Equipment cycles and revenue/protect equipment/spare fleet requirements;
- Crew schedules and train and engine crew requirements;
- Overnight and midday storage yard requirements;
- Fleet characteristics: type, consist length, top speed, passenger-carrying capacity; and
- Train performance calculations for each train type and stopping pattern.

"Protect equipment" refers to extra trains that are positioned at key locations to provide more rapid recovery from delays.

Once a deterministic service plan solution has been identified using the sketch-plan tools, the train schedule and rail infrastructure characteristics can be input into the simulation model, and the operating plan validated based on results of simulations that introduce stochastic variations into the analysis. This validation step will be undertaken for basic alternatives and at key intervals in the project. Network and service variations that pivot off of these basic alternatives can be analyzed with confidence using the sketch-plan tools with relatively short turnaround times. Standard operating plan outputs (for each discrete scenario, based primarily on detailed simulations), include:

- Operating timetable, including all revenue and non-revenue train movements, and indicating windows for maintenance-of-way and operation of freight traffic; for blended service scenarios, timetables and other simulation outputs will be generated for the transit or passenger line(s) and those elements of the conventional network feeding or otherwise interacting with those lines;
- String-line diagrams;
- Terminal track-occupancy diagrams;

- Operating performance characteristics, by train type and train, for the deterministic base case and any alternative delay or perturbation scenarios analyzed, including extent, location, and causes of delay;
- Equipment cycles and revenue/protect equipment/spare fleet requirements;
- Crew schedules and train and engine crew requirements;
- Overnight and midday storage yard utilization;
- Train performance calculations for each train type and stopping pattern; and
- Rolling stock consists, length, top speed, acceleration/deceleration profiles, and other relevant characteristics.

The output from the ridership forecasts, when these are available, will be used as feedback to the operations planning process to optimize the level of service for the forecast years for which projections are provided. The sketch-plan model contains a module that takes as input station-to-station projections of daily ridership, allocates ridership by time of day and among the rail services available at each time of day, and calculates average load factors for the various trains operating in the schedule. The model can be used to shift ridership among available trains to balance load factors (a coarse approximation of the effects of variable pricing techniques). The frequency and stopping patterns of train services then can be adjusted to match projected ridership and new service plan parameters furnished to the ridership estimation process. One additional service planning and ridership estimation iteration usually is required to provide a reasonably accurate balance between the level of rail service provided and the projected level of patronage. As the alternatives are narrowed through the study process, the number of iterations and the precision with which the ridership and service levels are “equilibrated” will be increased.

7. CALCULATION EXAMPLES

Exhibit 8-77
List of Calculation Examples

Example	Description
1	High-capacity heavy rail
2	Heavy rail line with junction
3	Heavy rail with long dwell
4	Light rail with single-track section
5	Commuter rail with limited train paths
6	AGT with short trains
7	AGT with off-line stations
8	Aerial ropeway

CALCULATION EXAMPLE 1: HIGH-CAPACITY HEAVY RAIL

The Situation

A transit agency is planning to build a heavy rail transit line and wants to determine the minimum train separation possible with a cab signaling system and with a variable safety distance moving-block signaling system.

The Questions

1. What is the minimum train separation (ignoring station dwell time and operating margin effects) with each type of signaling system?
2. What is the non-interference headway with typical dwells and operating margins?
3. What is the resultant line capacity for a new system with higher-quality loading standards?

The Facts

The transit agency is planning to use trains consisting of a maximum of eight 75-ft cars. Trains will operate at a maximum of 60 mi/h (88 ft/s) and will be traveling at 32 mi/h (47 ft/s) when entering stations if the cab signaling system is chosen, and at 34 mi/h (50 ft/s) if a moving-block system is selected. (Note that these station approach speeds are the optimal speeds to achieve minimum train separation. Solving for the optimal approach speed directly is not a simple task and is best done using a computer spreadsheet's solver or goal seek function to automate the iterative process that is required.) The distance from the front of a stopped train to the station exit block is 33 ft. Assume that there are no grades into or out of stations and that no civil speed restrictions limit approach speeds to sub-optimal levels.

Outline of Solution

To answer this question, two equations must be used, one for each signaling system type: Equation 8-1 for cab signaling and Equation 8-3 for moving block. Note that these equations provide allowances for grades and line voltage effects that have been removed as they are not required to answer this question. The values for all variables are summarized in Exhibit 8-78, with default values from Exhibit 8-28 used for input data not specified above.

Exhibit 8-78
Calculation Example 1: Input Data

Value	Term	Description
calculated	t_{cs}	train control separation
600 ft	L_t	length of the longest train
35 ft	d_{eb}	distance from front of stopped train to start of station exit block in meters
47 ft/s (cab) 50 ft/s (moving block)	v_a	station approach speed
88 ft/s	v_{max}	maximum line speed (88 ft/s = 60 mi/h)
75%	f_{br}	braking safety factor—worst-case service braking is f_{br}% of specified normal rate—typically 75%
1.2 (cab) 1 (moving block)	b	separation safety factor—equivalent to number of braking distances (surrogate for blocks) that separate trains
3.0 s	t_{os}	time for overspeed governor to operate on automatic systems—driver sighting and reaction times on manual systems
0.5 s	t_{jl}	time lost to braking jerk limitation
1.5 s	t_{br}	brake system reaction time
4.3 ft/s^2	a	initial service acceleration rate
4.3 ft/s^2	d	service deceleration rate
20.5 ft	P_e	positioning error—moving block only

Computational Steps

Step 1a: Determine the Maximum Load Point Station

Because this example addresses a rail line that does not yet exist, the dwell time selected in Step 1c should reflect conditions at a maximum load point station.

Step 1b: Determine the Control System's Minimum Train Separation

(1) *With Cab Signaling*

The relevant equation is Equation 8-1, modified to remove dwell, operating margin, voltage, and grade elements:

$$t_{cs} = \sqrt{\frac{2(L_t + d_{eb})}{a}} + \frac{L_t}{v_a} + \left(\frac{1}{f_{br}} + b\right)\left(\frac{v_a}{2d}\right) + \frac{at_{os}^2}{2v_a}\left(1 - \frac{v_a}{v_{max}}\right) + t_{os} + t_{jl} + t_{br}$$

$$t_{cs} = \sqrt{\frac{2(600 + 35)}{4.3}} + \frac{600}{47} + \left(\frac{1}{0.75} + 1.2\right)\left(\frac{47}{2 \times 4.3}\right) + \frac{(4.3)(3)^2}{2 \times 47}\left(1 - \frac{47}{88}\right) + 3 + 0.5 + 1.5$$

Compare the approach speed producing the minimum train separation with any speed restrictions on the station approach.

$$t_{cs} = 17.2 + 12.8 + (2.53)(5.47) + (0.412)(0.534) + 3 + 0.5 + 1.5$$

$$t_{cs} = 49.1\ s$$

(2) *With Moving-Block Signaling*

The relevant equation is Equation 8-3, modified to remove dwell, operating margin, voltage, and grade:

$$t_{cs} = \frac{L_t + P_e}{v_a} + \left(\frac{100}{f_{br}} + b\right)\left(\frac{v_a}{2d}\right) + \frac{at_{os}^2}{2v_a}\left(1 - \frac{v_a}{v_{max}}\right) + t_{os} + t_{jl} + t_{br}$$

$$t_{cs} = \frac{600 + 20.5}{50} + \left(\frac{100}{75} + 1\right)\left(\frac{50}{2 \times 4.3}\right) + \frac{(4.3)(3)^2}{2 \times 50}\left(1 - \frac{50}{88}\right) + 3 + 0.5 + 1.5$$

$$t_{cs} = 12.4 + 13.6 + (0.387)(0.432) + 3 + 0.5 + 1.5$$

$$t_{cs} = 31.2\ s$$

The net result is that the minimum train separation at stations (ignoring the effects of station dwells and an operating margin at this point) would be 49.1 s with a cab signaling system or 31.2 s with a variable safety distance moving-block system and automatic train operation.

Step 1c: Determine the Average Dwell Time at the Critical Station

Step 1c in Section 5 presented four possible methods for determining the controlling dwells. Method 2, *Using Existing Dwell Time Data*, is not applicable to a new system. The simplest option is to use Method 1, which recommends a range of dwell values from 35 to 45 s. If there are no indications of any single, very high-volume stations (where the more complicated dwell calculations should be used) then a median value of 40 s can be selected.

Step 1d: Select an Operating Margin

Section 5 suggests that the more operating margin that can be incorporated in the headway the better, with 20 to 25 s as the best guide. Here, 25 s is selected to provide better reliability.

Step 2: Determine the Minimum Headway Associated with the Right-of-Way Type

Step 2 primarily applies to light rail lines and therefore can be skipped for this heavy analysis.

Step 3: Determine the Limiting Junction Headway

As this will be a new line, it is assumed that it will be designed so that any junctions will not constrain capacity.

Step 4: Check Power Supply Constraints

As this will be a new line, it is assumed that the power system will be designed to accommodate the desired headway.

Step 5: Determine the Controlling Headway

In the absence of other constraints, the controlling headway is the sum of the minimum train separation time (Step 1b), average dwell time at the critical station (Step 1c), and the operating margin (Step 1d). For cab signaling, this headway is 49.1 + 40 + 25 = 114.1 s. For moving-block signaling, this headway is 31.2 + 40 + 25 = 96.2 s.

Step 6: Determine Terminal Layover Time

As this will be a new line, it is assumed that the terminals will be designed so as not to constrain capacity.

Step 7: Determine Train Throughput

Train throughput is 3,600 seconds per hour divided by the controlling headway. For cab signaling, this results in 31 trains per hour (rounded down). For moving-block signaling, this results in 37 trains per hour (rounded down).

Step 8: Determine Person Capacity

In the absence of a specific vehicle, the text accompanying Step 8 in Section 5 indicates that a recommended comfortable heavy rail car loading for a new system is 1.8 passengers per linear foot of train length, inclusive of diversity allowances. At this loading level, each specified train of eight 75-ft-long cars can carry 8 × 1.8 × 75 = 1,080 passengers.

The Results

Multiplying the number of passengers per train by the number of trains per hour provides passengers per peak hour direction per track of 33,480 p/h/dir and 39,960 p/h/dir, respectively. Reflecting the approximations used in this determination, the results should be rounded down to the nearest 1,000—33,000 and 39,000.

CALCULATION EXAMPLE 2: HEAVY RAIL LINE WITH JUNCTION

The Situation

The transit agency from Example 1 has decided to use a variable safety distance moving-block signaling system. The agency would now like to know if it can economize on construction by building a flat junction at a point where two of its lines will diverge. The agency's long-term plan is to run a 2-min headway through the junction, with service split equally between the two branches.

The Question

Can a flat junction on this proposed system support a 2-min headway or must a flying junction be constructed?

The Facts

Many of the variables are the same as those used in the previous example. In addition, the agency plans to build its tracks 16 ft (5 m) apart and use #10 turnouts (switches) with a throw-and-lock time of 6 s at mainline junctions. To make operations through a flat junction reliable, the agency plans to increase the operating margin to 45 s, hence the headway increases from 100 s (36 trains per hour) to 120 s (30 trains per hour).

Computational Steps

Equation 8-10 is used to estimate the capacity of a flat junction. The variables used in the equation are summarized in Exhibit 8-79, with default values from Exhibit 8-28 used for values not specified above.

Value	Term	Description
Calculated	h_j	limiting headway at junction
31.2 s	t_{cs}	line headway, from Calculation Example 1, Step 1b(2)
600 ft	L_t	train length
9.62	f_{sa}	switch angle factor (9.62 for a #10 switch, from Equation 8-10)
16 ft	d_{ts}	track separation
4.3 ft/s^2	a	initial service acceleration rate
4.3 ft/s^2	d	service deceleration rate
88 ft/s	v_{max}	maximum line speed (88 ft/s = 60 mi/h)
6 s	t_{sw}	switch throw and lock time
45 s	t_{om}	operating margin

Exhibit 8-79
Calculation Example 2: Input Data

Substituting the known variables into the equation produces:

$$h_j = t_{cs} + \sqrt{\frac{2(L_t + 2f_{sa}d)}{a}} + \frac{v_{max}}{a+d} + t_{sw} + t_{om}$$

$$h_j = 31.2 + \sqrt{\frac{2(600 + [2 \times 9.62 \times 4.3])}{4.3}} + \frac{88}{4.3 + 4.3} + 6 + 45$$

$$h_j = 31.2 + 17.8 + 10.2 + 6 + 45$$

$$h_j = 110.2 \text{ s}$$

The Results

While the resulting value of h_j would appear to support 2-min headways, it is about 10 s less than the planned headway. Based on this narrow margin, it would be prudent to opt for a flying junction rather than risk service disruptions with a flat junction—even with the operating margin increased to 45 s. This is consistent with the recommendation in Section 5 that junctions should be grade separated at headways below 3 min.

CALCULATION EXAMPLE 3: HEAVY RAIL WITH LONG DWELL

The Situation

A busy heavy rail line operates through a major transfer station with long station dwell times.

The Question

What is the maximum person capacity through this station?

The Facts

- A generous loading standard means more passengers seated.
- The transit agency's loading standard is 1.8 passengers per linear foot of train length during the peak 15 min.
- Service is provided by ten-car trains with each car being 75 ft long.
- The dwell time at this station averages 30 s with a standard deviation of 21 s.

- There is a 1.5% downgrade into the station and a 1.5% upgrade out of the station.
- The line is automated and uses moving-block signaling.
- Train operators are responsible for closing the doors and initiating acceleration; this delay is incorporated into the dwell time.
- Trains are evenly loaded over their length.

Outline of Solution

The solution consists of three key steps: (a) determining each train's passenger capacity, (b) determining the minimum train separation based on the signaling system and train length, and (c) incorporating the station dwell time and an operating margin. To determine the non-interference headway, allowances for dwell time and an operating margin must be added to the minimum train separation time. The results of these steps can then be combined to produce the line capacity based on the parameters given.

Computational Steps

Step 1a: Determine the Maximum Load Point Station

The selected station is the maximum load point station.

Step 1b: Determine the Control System's Minimum Train Separation

This step requires the use of Equation 8-3. The values for all variables are summarized in Exhibit 8-80, with default values from Exhibit 8-28 used for values not specified above.

Exhibit 8-80
Calculation Example 3: Input Data

Value	Term	Description
calculated	t_{cs}	train control separation
750 ft	L_t	length of the longest train
35 ft	d_{eb}	distance from front of stopped train to start of station exit block in meters
50 ft/s	v_a	station approach speed (50 ft/s = 34 mi/h)
88 ft/s	v_{max}	maximum line speed (88 ft/s = 60 mi/h)
75%	f_{br}	braking safety factor—worst-case service braking is f_{br}% of specified normal rate—typically 75%
1	b	separation safety factor—equivalent to number of braking distances (surrogate for blocks) that separate trains
3.0 s	t_{os}	time for overspeed governor to operate on automatic systems—driver sighting and reaction times on manual systems
0.5 s	t_{jl}	time lost to braking jerk limitation
1.5 s	t_{br}	brake system reaction time
4.3 ft/s^2	a	initial service acceleration rate
4.3 ft/s^2	d	service deceleration rate
32 ft/s^2	a_g	acceleration due to gravity
-1.5%	G_i	grade into the station
+1.5%	G_o	grade out of the station
90%	l_v	line voltage as percentage of specification
20.5 ft	P_e	positioning error

Substituting the variables into Equation 8-3 produces:

$$t_{cs} = \frac{L_t + P_e}{v_a} + \left(\frac{1}{f_{br}} + b\right)\left(\frac{v_a}{2(d + a_g G_i)}\right) + \frac{(a + a_g G_o) l_v^2 t_{os}^2}{2 v_a}\left(1 - \frac{v_a}{v_{max}}\right) + t_{os} + t_{jl} + t_{br}$$

$$t_{cs} = \frac{750 + 20.5}{50} + \left(\frac{1}{0.75} + 1\right)\left(\frac{50}{2(4.3 + [32 \times -0.015])}\right) + \frac{(4.3 + [32 \times -0.015])(0.9)^2(3.0)^2}{2 \times 50}\left(1 - \frac{50}{88}\right) + 3.0 + 0.5 + 1.5$$

$$t_{cs} = 15.4 + (2.33)(6.54) + (0.278)(0.432) + 3.0 + 0.5 + 1.5$$

$$t_{cs} = 35.6 \text{ s}$$

Step 1c: Determine the Average Dwell Time at the Critical Station

From the facts provided, based on field measurements, the average dwell time at this station is 30 s.

Step 1d: Select an Operating Margin

As this is an existing line, the operating margin can be estimated as two times the standard deviation of dwell times—in this case 42 s, based on the facts provided.

Steps 2, 3, 4, and 6

In the absence of other information, it is assumed that junctions, turnbacks, power system, etc. do not constrain line capacity.

Step 5: Determine the Controlling Headway

The controlling headway is the sum of the minimum train separation time (36 s), average dwell time at the critical station (30 s), and the operating margin (42 s), or 108 s.

Step 7: Determine Train Throughput

Train throughput is 3,600 seconds per hour divided by the controlling headway (108 s), resulting in 33 trains per hour (rounded down).

Step 8: Determine Person Capacity

This step is straightforward and is based on the number of cars in each train, the length of each car, and the number of passenger spaces per unit of car length. Because the agency's loading standard is based on peak 15-min conditions, a peak-hour factor must be used. In the absence of other information, a PHF of 0.80 is suggested in Section 5 for heavy rail. The PHF accounts for lower passenger demand during the other 45 min of the peak hour, which results in unused capacity. If the agency policy had been to maintain an average loading of 1.8 p/ft length throughout the peak hour, resulting in more crowded peak 15-min conditions, no PHF would have been needed, as the 1.8 p/ft length value already incorporates peak-hour loading diversity.

When a PHF is and is not needed.

(10 cars/train)(75 ft/car)(1.8 p/ft)(0.80) = 1,080 p/train

Multiplying 1,080 p/train by 33 train/h gives a person capacity of approximately 35,000 passengers during the peak hour in the peak direction through this station (rounded down to the nearest 1,000 to reflect the approximations used).

CALCULATION EXAMPLE 4: LIGHT RAIL WITH SINGLE-TRACK SECTION

The Situation

A light rail line operates with a single-track section.

The Question

What is the maximum possible service frequency?

The Facts

- Service is provided by three-car trains, with each car 85 ft long.
- The single track section is 4,000 ft long with one intermediate station, with an average dwell time of 20 s.
- The section is on a road with a speed limit of 30 mi/h.

Assumptions

- It is assumed that there are no other longer single-track sections on the line, nor any more restrictive limitations imposed elsewhere along the line.

Outline of Solution

The maximum possible service frequency is twice the travel time through the single-track section, plus an allowance for operational irregularities.

Computational Steps

The travel time over the single-track section can be calculated using Equation 8-5. The values for all variables are summarized in Exhibit 8-81, with default values from Exhibit 8-38 used for values not specified above:

Exhibit 8-81
Calculation Example 4: Input Data

Value	Term	Description
calculated	t_{st}	time to cover single-track section
1.1	S_m	speed margin
1	N_{st}	number of stations on the single-track section
44.0 ft/s	$v_{max,st}$	maximum speed in single-track section (44.0 ft/s = 30 mi/h)
0.5 s	t_{jl}	time lost to braking jerk limitation
4.3 ft/s^2	d	deceleration rate
1.5 s	t_{br}	brake system reaction time
4,000 ft	L_{st}	length of single-track section
255 ft	L_t	train length
20 s	t_d	dwell time
6.0 s	t_{sw}	switch throw-and-lock time
20 s	t_{om}	operating margin (middle of range from Exhibit 8-38)

Substituting these results into Equation 8-5 produces:

$$t_{st} = S_m \left[\left(\frac{N_{st} + 1}{2} \right) \left(\frac{3v_{max}}{d} + t_{jl} + t_{br} \right) + \frac{L_{st} + L_t}{v_{max,st}} \right] + N_{st}t_d + t_{sw} + t_{om}$$

$$t_{st} = 1.1 \left[\left(\frac{1+1}{2} \right) \left(\frac{3 \times 44}{4.3} + 0.5 + 1.5 \right) + \frac{4{,}000 + 255}{44} \right] + (1 \times 20) + 6 + 20$$

$$t_{st} = 1.1[(1)(30.7 + 0.5 + 1.5) + 96.7] + 20 + 6 + 20$$

$$t_{st} = 188\ s$$

The minimum headway that can be operated over the single-track section is twice this time, or 376 s. Normally, this headway would be rounded up to the nearest even hourly headway of 480 s (7½ min), resulting in 8 trains per hour per direction.

A 7½-min headway is often expressed in transit timetables as service every 7 to 8 min.

The Results

The single-track section can support up to 8 trains per hour per direction. However, if there is significant on-street running elsewhere on the line, it is unlikely that service can be maintained with sufficient regularity that trains will not be held up at the entrance to the single-track section, waiting for the opposing train to clear. In this case, it would be prudent to increase the minimum headway to the next even interval, or trains every 10 min.

Comments

Single-track working in emergencies.

In the event of track maintenance or an emergency such as a traffic accident, failed train, or derailment, crossovers are usually provided to permit single-track working around the obstruction. For long-term obstructions—such as a track renewal program—temporary crossovers, called *shoo flys*, can be used. Where a signaling system is used, this operation is only possible if either (a) the signaling system is equipped for two-way operation on either track, or (b) operations are reverted to a slower manual, line-of-sight operation. Such emergency operation is then limited to a frequency as calculated by Equation 8-5 and line capacity is reduced.

As an example, if normal service on a double-track line is a train every 5 min, but 4,000 ft of single-track operation is needed to pass an obstruction, service will be limited to 7½ minutes. Nominal capacity will be reduced from 12 to 8 trains per hour (i.e., by one-third). This reduction is sufficiently small that it may be accommodated temporarily by accepting higher levels of crowding. Passengers are generally willing to accept this in emergency conditions.

Capacity reduction with single track working.

Longer single-track sections will reduce capacity further. This loss may be made up where operational policies and signaling systems permit platooning trains over the single-track section. Two or three trains can follow each other closely under line-of-sight operating practice at lower speeds. Full capacity may be restored, but additional trains and drivers will be required to compensate for the slower speeds and waiting time while trains accumulate to form a platoon. In addition, passengers will experience long waits between train platoons relative to the normal headway, which will result in more crowded station platforms.

Wrong-side or wrong-way working over line sections with grade crossings on on-street track can be confusing to motorists and pedestrians and can be hazardous. As a result, many light rail operators prohibit such operations except where there are no alternatives, such as in tunnels or subways. Instead, their emergency planning calls for a bus bridge around any blockage that is expected to take a significant time to clear. All North America light rail operators also have or are affiliated with major bus operations and can expect to obtain buses and drivers for such emergency use on short notice—usually by scavenging buses from nearby high-frequency routes.

CALCULATION EXAMPLE 5: COMMUTER RAIL WITH LIMITED TRAIN PATHS

The Situation

An existing commuter rail agency would like to expand its operations to a new route that is owned by a freight railroad.

The Question

Based on the constraints given below, can the commuter rail agency provide service on the new line with its current single-level car fleet, or must it order new double-level cars for the line?

The Facts

- The freight railroad will only allow six commuter rail trains per hour to use its line.
- Physical constraints mean that station platforms on the new line can be no more than eight cars in length.
- The commuter rail agency currently uses single-level cars that have 120 seats but is considering the purchase of two-level cars with 180 seats, although it would prefer to purchase more single-level cars to maintain a standard fleet.
- The agency has a policy of planning service based on cars being at 90% of seated capacity.
- The agency would like to be able to accommodate a flow of 6,000 passengers per hour in the peak hour.
- Train scheduling can be adjusted to meet the peak 15-min demand, provided no more than six trains are operated per hour.
- Trains are limited by railroad contract, but they can be spaced through the peak hour to best match demand.

Outline of Solution

To determine which car type, if either, can satisfy the agency's capacity needs, the hourly capacity of the line using each car type must be determined. In this example, this procedure is simplified by the agency's ability to schedule trains to meet the peak 15-min demand, avoiding the need to consider the temporal distribution of travel. The capacity that can be provided with each car type should be considered independently.

Computational Steps

The hourly capacity P is determined as follows:

$$P = (\text{passengers per car}) \times (\text{cars per train}) \times (\text{trains per hour}) \times (\text{PHF})$$

Single-level cars

The effective capacity per car is 90% of 120, or 108 passengers. An eight-car train of single-level cars could thus carry 864 passengers. With six trains per hour, the capacity is 5,184 passengers per hour, rounded down to 5,100 to reflect the approximations used.

Two-level cars

The effective capacity per car is 90% of 180, or 162 passengers. An eight-car train of two-level cars could thus carry 1,296 passengers. With six trains per hour, the capacity is 7,776 passengers per hour, rounded down to 7,700 to reflect the approximations used.

The Results

Since eight-car trains of single-level cars are unable to handle the predicted demand of 6,000 passengers per hour, it appears that the agency should plan on ordering two-level cars for use on this route. The calculation above shows that the two-level cars can accommodate the projected demand with some room for ridership growth.

The only alternative to purchasing the two-level cars would be to operate longer trains and assign passengers to cars according to their destination station, since not all cars would be adjacent to a platform at all stations. This would only work if the platforms at major terminal stations could accommodate all the cars of each train. As it complicates train operations and would likely create passenger confusion, the option of purchasing two-level cars is preferable.

Longer trains that overhang platforms are a poor compromise.

CALCULATION EXAMPLE 6: AGT WITH SHORT TRAINS

The Situation

An automated feeder line is planned from a new suburban office development to an existing heavy rail station.

The Question

Based on the use of advanced train control systems, what is the design capacity of this line?

The Facts

The developer wants to incorporate the AGT stations in an elevator lobby on the second floor of each building, which limits station length to 85 ft. AGT trains used in institutional settings normally require most users to stand; the developer would like to provide a comfortable standing environment for the short trip. Most AGT systems are proprietary and the manufacturer would provide capacity capabilities. In this case, the developer does not wish to approach a manufacturer at this stage. However, a review of similar AGT systems suggests that the vehicle would be at least 8 ft wide.

Computational Steps

Exhibit 8-50 shows that an AGT moving-block train control system can provide a minimum train separation of 13.4 s with 80-ft trains. Adding relatively high dwell time and operating margin values of 40 and 25 s respectively would result in a minimum headway of 78.4 s, which would normally be rounded up to 80 s to provide an integer number of trains per hour. Section 5 states that headways shorter than 90 s are possible but may limit dwell times and constrain the operating margin. They should be considered with caution unless off-line stations are adopted. As the developer has indicated a relatively relaxed loading level, it is realistic to expect dwells to be lower than normal and hence the 80-s headway can be accepted as practical. This equates to 3,600/80 or 45 trains per hour.

The 85-ft platform can hold two 40-ft cars, a common AGT size and comparable to a transit bus. As no specific car design is available at this point, it is reasonable to assume that each car will have a total interior area of at least 288 ft^2 (40-ft length by at least 8-ft width, with each dimension reduced by 8 in. to account for wall thickness). Using a comfortable loading standard of 5.4 ft^2 per passenger gives a total car design passenger capacity of at least 53. Note that an AGT car of this size, on a short-distance line, would normally be rated for 100 passengers, packed closer together.

The resultant design capacity at the preferred loading level is the number of trains per hour, 45, multiplied by the number of passengers per car, 53, and the number of cars per train, 2, or 4,770 passengers per peak-hour direction. As always with such calculations where there are approximations, the number should be rounded down, in this case to 4,700 p/h/dir. Note that this result would need to be multiplied by a peak-hour factor to determine the number of people that could be accommodated without exceeding the preferred loading level at any point during the hour.

Comments

Because the proposed AGT line would connect to a heavy rail station, an 80-s headway might be unnecessary. In operation, a headway corresponding to either the heavy rail headway or half the heavy rail headway (depending on the proportion of AGT ridership expected to come from each direction) might be used, with the goal to serve all passengers on an arriving heavy rail train on the next departing AGT train.

CALCULATION EXAMPLE 7: AGT WITH OFF-LINE STATIONS

The Situation

The developer from Example 6 is expanding the suburban office development to include a major shopping complex and recreation facility with an ice hockey arena.

The Question

How can the AGT line be expanded to handle this load?

The Facts

- Ridership estimates are that the system will handle one-third of the arena's maximum capacity of 24,000 people, plus an estimated demand of 1,200 passengers per hour from the shopping complex.
- Two adjacent stations serve the sports arena, while the shopping center has three stations.

- The developer has contracted with the office building tenants to run trains at least every 6 min until midnight each day, including weekends and holidays.

Computational Steps

A peak-hour factor is not applied to special event person capacity calculations, as the offered capacity is generally fully utilized and passengers do not expect to be able to board the first train that arrives.

To handle 24,000/3 + 1,200 = 9,200 passengers per hour, one solution would be to operate longer trains with higher occupancy and to omit stops in the office buildings with their short stations. The 45-train-per-hour capacity is no longer practical as heavy loads at the two sports arena stations will extend dwells and longer trains will increase the minimum train separation. The capacity is decreased to 40 trains per hour, or 90-s headways. Ten of these trains, one every 6 min, will remain short to serve the office complex. However, during events at the arena, passengers will tolerate a higher loading level. At a maximum schedule load of 3.2 ft^2 per passenger, each car can accommodate at least 90 passengers, or 180 per short train of two cars.

The ten short trains can serve 1,800 passengers. The remaining 30 trains must carry 7,400 passengers per hour. This results in 7,400/30 or 247 passengers per train. Three-car trains holding 90 p/car would be required to meet the demand.

Section 5 stated that off-line stations permit headways that are partly independent of station dwell time with throughput that of the control system minimum train separation, plus an allowance for switch operation, lock and clearance, and a reduced operating margin. Exhibit 8-50 shows that a moving-block signaling system with 80-ft trains has a minimum train separation of 13.4 s. Allowing an operating allowance for merging trains of 45 s and rounding up results in permitted headways as low as 60 s, or 60 trains per hour. In this case, the demand of 9,200 passengers per hour with 180 passengers per train can be met by 52 trains, within the 60-train maximum.

It is not always economical to meet occasional peak demands with rail transit.

Off-line stations would permit trains to operate directly from each arena station to the heavy rail station. However, economics enter the picture. It is unlikely that the developer would be willing to build more expensive off-line stations and purchase addition rolling stock for a sports arena demand that only occurs a few days a year. It is more likely that the system would be designed for maximum office and shopping complex demands. When a sports event takes place, the AGT line would be filled to capacity and the overload would be handled by transit authority buses—of which there is a surplus at the off-peak hours typical of sport event starts and finishes.

CALCULATION EXAMPLE 8: AERIAL ROPEWAY

The Situation

A university hospital is located on a bluff above a river. The university has run out of room to expand on the bluff and is seeking to move some of its operations to a new campus along the riverfront. For the two campuses to function efficiently as a single entity, good transportation links will need to be provided between them. The university is exploring various means to provide these links, including shuttle buses and roadway and parking improvements. Another option under consideration is a direct link between the two campuses using an aerial ropeway, either an aerial tramway or a detachable-grip aerial lift (gondola) system.

The Questions

1. For the aerial tramway, how large will the carriers need to be to handle the projected passenger demand?
2. For the gondola, how many carriers will be needed?

The Facts

Based on the university functions to be located on the riverfront and an estimate of total faculty, staff, and student sizes at build-out, the university estimates that a total of 750 persons will need to be carried in the peak direction during the peak hour. The line would be approximately 2,600 ft long, with no intermediate stations. A decision on a specific manufacturer has not been made; however, as a starting point, assume that the aerial tramway cabin door would be wide enough that three people can walk through at a time and that the gondola carriers would seat eight people each.

Aerial tramway dwell time includes the time to unload and load passengers from the cabin, plus an assumed allowance of 60 s to (a) clear exiting passengers from the platform and (b) perform safety and communications checks prior to the carrier departing. Maximum acceleration and deceleration is 0.65 ft/s^2.

Gondola carriers take 60 s to traverse each station after detaching from the line. The carriers move at creep speed (0.8 ft/s) through the station to allow passenger loading and unloading.

Outline of Solution

Aerial Tramway

Aerial tramway capacity is based on the number of carriers used (one or two, two is typical), the number of stops per direction (one, in this case), station dwell time (not yet known), line length (given), line speed (a user decision), and the size of the carriers (a user decision). Passenger service time will be based on the time to clear a full cabin, and then load a full cabin. Several combinations of line speeds and cabin sizes may need to be tried in developing a solution.

Gondola

Gondola capacity is based on the spacing between carriers (not yet known) and the average line speed (a user decision). To solve this problem, the minimum number of carriers needed to serve the demand will be calculated by working backward from the required capacity.

Computational Steps

Aerial Tramway

As a starting point, a 60-passenger cabin and the fastest possible line speed (39 ft/s) will be selected. At a maximum acceleration rate of 0.65 ft/s^2, it takes 60 s (39 divided by 0.65) to reach line speed. The average speed during acceleration is half the line speed, or 19.5 ft/s. As a result, the carrier would travel 1,170 ft (60 s multiplied by 19.5 ft/s) during acceleration. The carrier would travel another 1,170 ft during deceleration, meaning that it would only travel 260 ft at line speed (2,600-ft line length, minus two times 1,170 ft). The total trip time would be about 127 s.

Spending only 10% of the trip length at line speed would be inefficient, so a lower line speed should be tried. At a line speed of 30 ft/s, acceleration and deceleration would take 46 s each and would cover a total distance of 1,380 ft. As a result, the carrier could travel at line speed for 1,220 ft (nearly half the distance) and could cover that distance in 41 s. The total trip time would be about 133 s. The corresponding average line speed would be 2,600 ft divided by 133 s, or 19.5 ft/s.

In the absence of other data, Exhibit 8-13 can be consulted to determine average passenger boarding and alighting times for a high-volume level doorway. From a review of the data provided in the exhibit, 1.8 s per alighting passenger per door channel and 1.9 seconds per boarding passenger per door channel can be chosen as median values. With three door channels, it takes 36 s on average for passengers to exit a full 60-passenger cabin (60 p times 1.8 s/p, divided by 3 door channels), and 38 s to board. The total dwell time, including the 60-s allowance for various checks discussed in the facts of the problem, is 134 s.

All the information needed to calculate line capacity is now known. Entering this information into Equation 8-16 gives:

$$T = \frac{1{,}800N_v}{(N_s t_d) + \frac{L_l}{v_l}} = \frac{(1{,}800 \text{ s/h})(2 \text{ veh})}{(1)(134 \text{ s}) + \frac{(2{,}600 \text{ ft})}{(19.5 \text{ ft/s})}} = \frac{3{,}600 \text{ veh-s/h}}{267 \text{ s}} = 13 \text{ veh/h}$$

Multiplying 13 carriers per hour by 60 passengers per carrier gives a theoretical directional capacity of 780 passengers per hour, which is more than the required 750 passengers per hour. However, because passengers are not likely to arrive evenly throughout the hour, a peak-hour factor should be applied. Using a PHF of 0.90, the directional design person capacity of the system is about 700 passengers per hour, which is insufficient to avoid pass-ups.

Repeating the above process with an 80-passenger cabin results in a 159-s dwell time, with all other input values remaining the same. The resulting line capacity is 12 carriers per hour, which provides a directional design person capacity of about 865 passengers per hour when a peak-hour factor of 0.90 is applied.

Gondola

Since the only thing known about the gondola system is an assumed carrier size (eight passengers), the number of carriers required will be determined by working backward from the required capacity. Using a PHF of 0.90, a maximum directional capacity of 833 passengers per hour is needed (demand of 750 passengers per hour, divided by 0.90). Dividing this capacity by eight passengers per carrier results in 105 eight-passenger carrier arrivals per hour required at a station. However, because each carrier will make more than one trip each hour, the number of actual carriers required will be smaller.

A carrier traveling at a line speed of 20 ft/s (the maximum for a detachable-grip lift) requires 130 s to travel the length of the line. Therefore, a round trip on the line takes 260 s. In addition, the carriers take 1 min to travel through each station at creep speed, adding another 120 s to the round-trip journey. Consequently, a carrier makes one round trip every 380 s, or 9.47 round trips per hour. The number of carriers that will provide the required number of hourly station arrivals is 105 arrivals per hour divided by 9.47 arrivals per carrier per hour, or 12 carriers (rounded up).

The Results

Although the aerial tramway carrier travels twice as fast as a gondola at their respective maximum line speeds, it takes much longer to accelerate and decelerate the aerial tramway carrier. As it turned out, the travel times of the two modes were nearly the same over the length of the relatively short route. The headway between aerial tramway carriers is approximately 5 min, while the headway between gondolas is about 32 s.

8. REFERENCES

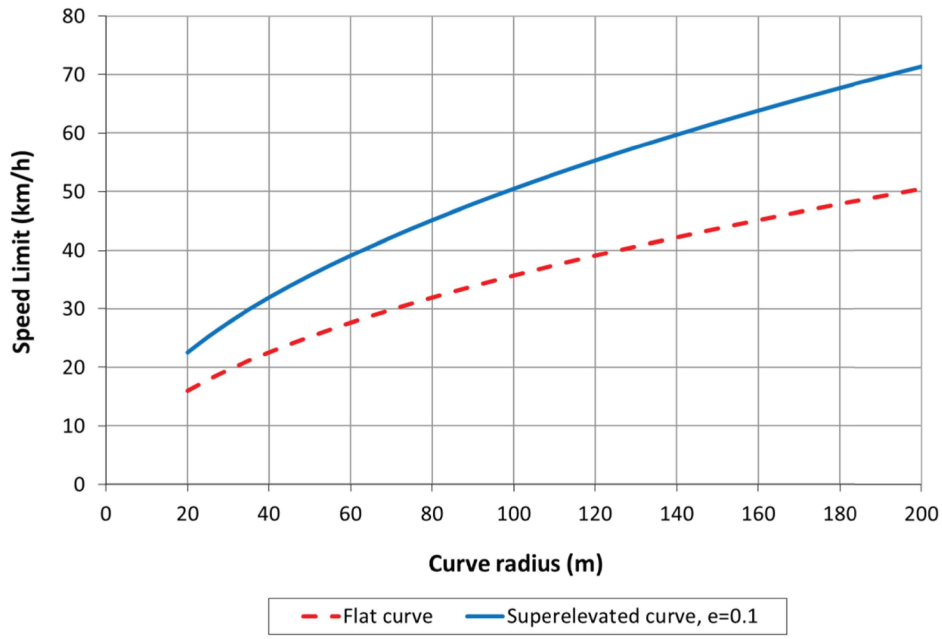
Links to the TCRP reports listed here can be found on the accompanying CD-ROM.

1. Parkinson, T. and I. Fisher. *TCRP Report 13: Rail Transit Capacity.* Transportation Research Board, National Research Council, Washington, D.C., 1996. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_13-a.pdf
2. American Public Transportation Association. *Public Transportation Vehicle Database.* Table 22: U.S. Average New Vehicle Costs for 2010 and 2011 Vehicles by Type. Washington, D.C. http://www.apta.com/resources/statistics/Documents/table22_vehcosttransitlength2011.pdf, accessed July 3, 2012.
3. Dure, D. *Evaluating and Preventing Capacity Loss when Designing Train Control to Enforce NFPA 130 Compliance.* Presented at the 2010 APTA Rail Conference, Vancouver, Canada, June 2010. http://www.apta.com/mc/rail/previous/2010/Papers/Evaluating-and-Preventing-Capacity-Loss.pdf
4. Korve, H.W., J.I. Farran, D.M. Mansel, H.S. Levinson, T. Chira-Chivala, and D.R. Ragland. *TCRP Report 17: Integration of Light Rail Transit into City Streets.* Transportation Research Board, National Research Council, Washington, D.C., 1996. http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_17-a.pdf
5. Kittelson & Associates, Inc.; KFH Group, Inc.; Parsons Brinckerhoff Quade and Douglass, Inc.; and K. Hunter-Zaworski. *TCRP Report 100: Transit Capacity and Quality of Service Manual, 2nd Edition.* Transportation Research Board of the National Academies, Washington, D.C., 2003. http://www.trb.org/Publications/Blurbs/153590.aspx
6. National Fire Protection Association. *NFPA 130 Standard for Fixed Guideway Transit and Passenger Rail Systems.* Washington, D.C., 2010.
7. Tereschuck, P.D. San Diego Light Rail Transit Goes to the Superbowl. Presented at 79th Annual Meeting of the Transportation Research Board., Washington, D.C., January 2000.
8. American Railway Engineering and Maintenance-of-Way Association. *AREMA Manual for Railway Engineering.* Landover, Md., 2000.
9. Allen, D.W. Practical Limits of Single-Track Light Rail Transit Operation. In *Transportation Research Record 1361.* Transportation Research Board, National Research Council, Washington, D.C., 1992.
10. Levinson, H.S. Capacity Concepts for Street-Running Light Rail Transit. In *Proceedings of the Second International Symposium on Highway Capacity.* Australian Road Research Board Ltd. and Transportation Research Board, Sydney, Australia, August 1994.
11. Wilson, Rick. Single Track Capacity Issues, Part 1. Presented at the Workshop on Railroad Capacity and Corridor Planning at 81st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2002.
12. Galloway, D. Train Performance Characteristics. Presented at the Workshop on Railroad Capacity and Corridor Planning at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2002.

13. Schaefer, B. Introduction. Presented at the Workshop on Railroad Capacity and Corridor Planning at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2002.
14. Reistrup, P. Prioritizing Train Movements. Presented at the Workshop on Railroad Capacity and Corridor Planning at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2002.
15. Reistrup, P. Summary. Presented at the Workshop on Railroad Capacity and Corridor Planning at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2002.
16. Auer, J.H. Rail-Transit People-Mover Headway Comparison. In *IEEE Transactions on Vehicular Technology,* Institute of Electrical and Electronics Engineers, 1974.
17. Schneigert, Z. *Aerial Tramways and Funicular Railways.* Pergamon Press, Oxford, U.K., 1966.
18. Bing, A.J., E.W. Beshers, M. Chavez, D.P. Simpson, E.S. Horowitz, and W.E. Zullig, Jr. *NCHRP Report 657: Guidebook for Implementing Passenger Rail Service on Shared Passenger and Freight Corridors.* Transportation Research Board of the National Academies, Washington, D.C., 2010.

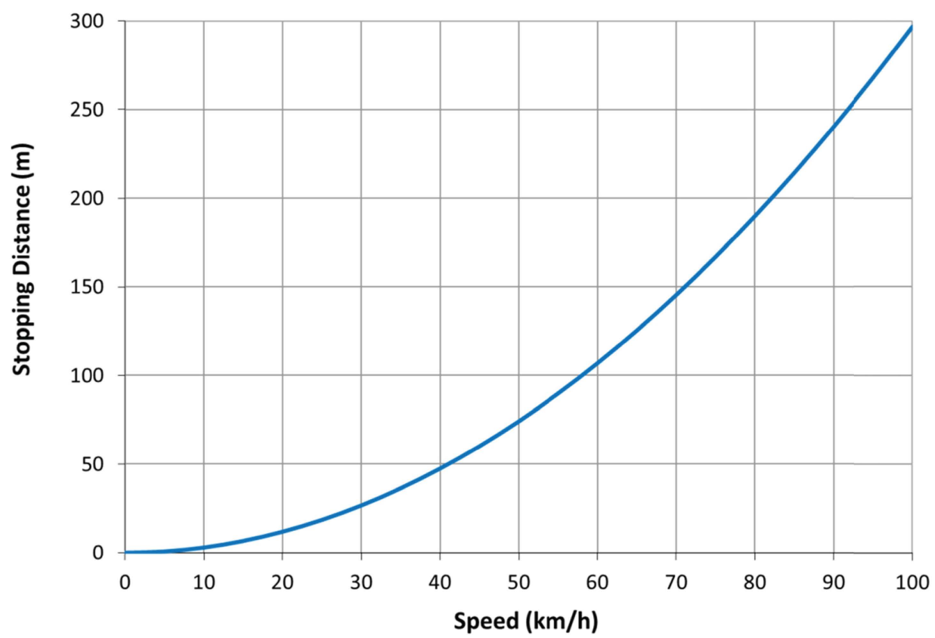
APPENDIX A: EXHIBITS IN METRIC UNITS

Exhibit 8-31m
Maximum Speed Limits on Curves



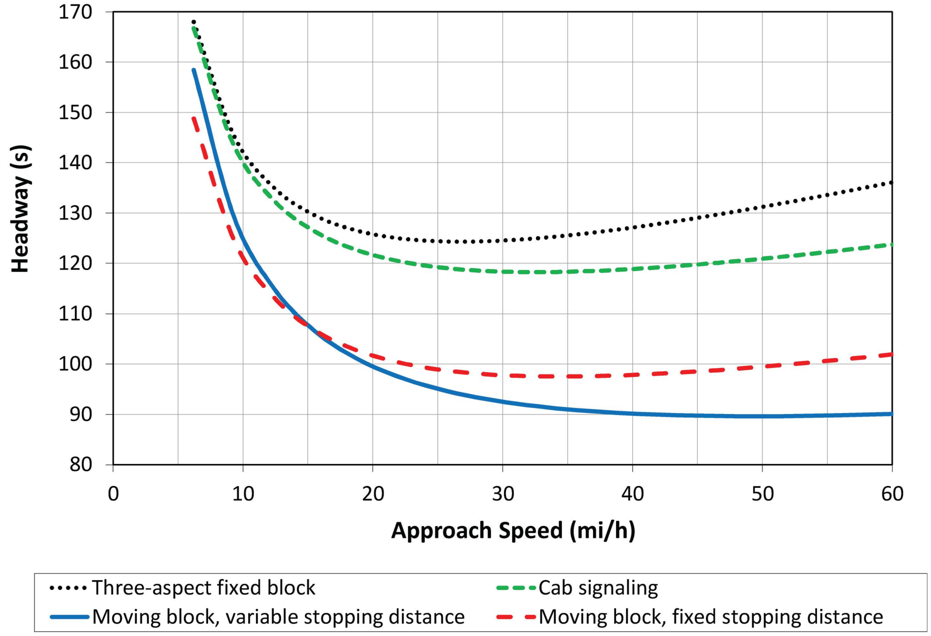
Source: *TCRP Report 13* (*1*).
Note: Transition spirals are not taken into account.

Exhibit 8-33m
Typical Stopping Distance as a Function of Speed



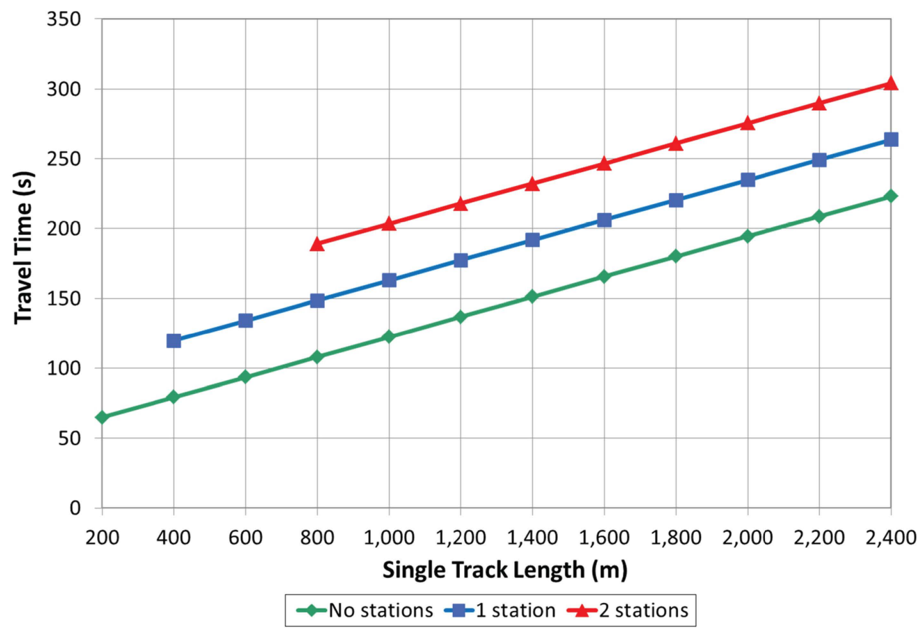
Source: *TCRP Report 13* (*1*).

Exhibit 8-36m
Typical Moving-Block Station Headways Compared with Conventional Fixed-Block Systems



Source: *TCRP Report 13* (*1*).

Exhibit 8-39m
Light Rail Travel Time Over Single-Track Section



Source: *TCRP Report 13* (*1*).
Note: Assumes speed limit of 55 km/h, train length of 55 m, 20-s dwell time, 20-s operating margin, and other data as per Exhibit 8-38. *The recommended closest headway is twice this time.*

Exhibit 8-49m
Average Commuter Rail Operating Speeds

Station Spacing (km)	Average Operating Speed (km/h) P/W = 3.0	P/W = 5.8	P/W = 9.1
Average Dwell Time = 30 s			
1.6	27.0	32.7	35.9
3.2	41.5	49.7	56.4
6.4	58.6	71.0	78.2
8.0	64.9	78.4	84.8
Average Dwell Time = 60 s			
1.6	23.8	28.0	30.3
3.2	37.5	44.1	49.3
6.4	54.4	65.0	71.0
8.0	60.9	72.5	78.1

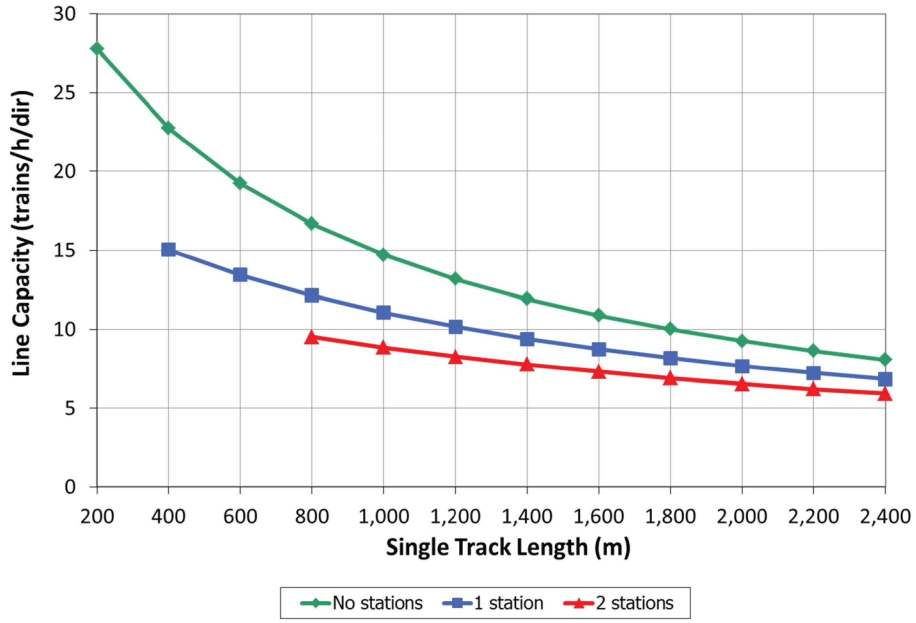
Source: Galloway (*12*).
Note: P/W = power-to-weight ratio. Assumes 127-km/h speed limit, no grades, and no delays due to other trains.

Exhibit 8-51m
Suggested AGT Separation Calculation Default Values

Default Value General Method	AGT	Term	Description
6.25 m	6.25 m	P_e	positioning error
200 m	**50 m**	L	**length of the longest train**
10 m	**0 m**	d_{eb}	**distance from front of train to exit block**
75%	75%	f_{br}	% service braking rate
2.4	**4**	b	**train detection uncertainty constant — fixed block**
1	1	b	train detection uncertainty constant — moving block
3 s	**1 s**	t_{os}	**time for overspeed governor to operate**
0.5 s	0.5 s	t_{jl}	time lost to braking jerk limitation
1.3 m/s^2	**0.6 m/s^2**	a	**service acceleration rate**
1.3 m/s^2	**1.0 m/s^2**	d	**service deceleration rate**
1.5 s	**0.5 s**	t_{br}	**brake system reaction time**
100 km/h	**80 km/h**	v_{max}	**maximum line velocity**
50 m	**25 m**	S_{mb}	**moving-block safety distance**

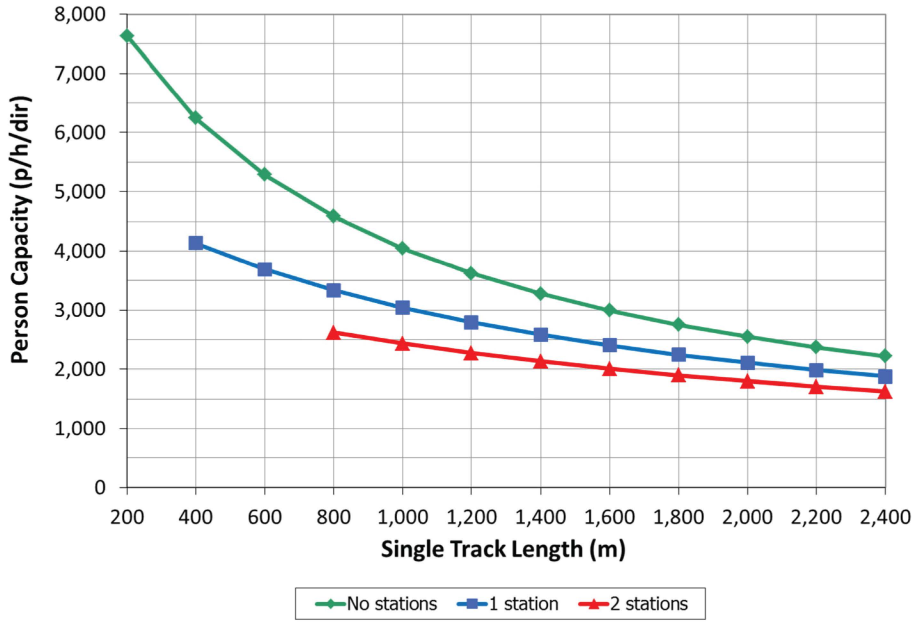
Source: *TCRP Report 13* (*1*).
Note: Bold type indicates AGT default values that differ from other rail transit values.

Exhibit 8-59m
Single-Track Line Capacity—Two-Car Light Rail Trains



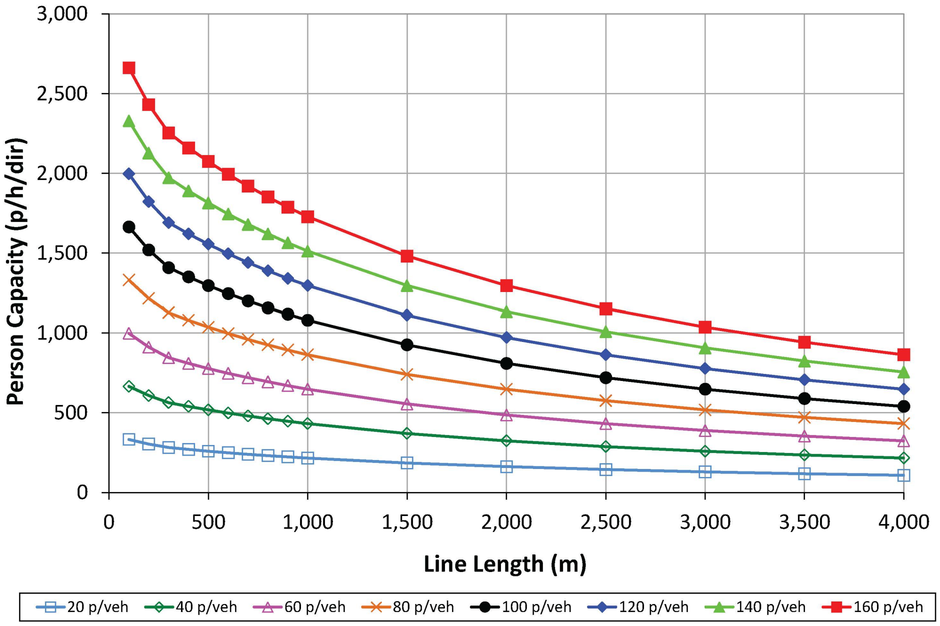
Source: *TCRP Report 13* (*1*).
Note: Assumes 55-km/h speed limit, 55-m train length, 20-s dwell time, and 20-s operating margin.

Exhibit 8-60m
Single-Track Person Capacity—Two-Car Light Rail Trains



Source: *TCRP Report 13* (*1*).
Note: Assumes 55-km/h speed limit, 55-m train length, 20-s dwell time, and 20-s operating margin.

Exhibit 8-63m
Reversible Ropeway Person Capacity



Note: Assumes 10-m/s line speed, 0.2-m/s^2 acceleration, two-vehicle operation, no intermediate stations, 150-s dwell time, and 0.90 PHF.