

Bulk Service Queueing for Transit Resilience under Short Random Service Suspensions

Baichuan Mo¹, Li Jin^{2,3,*}, Zhengzhong Ricky You¹, Zuo-Jun Max Shen⁴,
Haris N. Koutsopoulos⁵, Jinhua Zhao⁶

¹Department of Civil Engineering, Tsinghua University, Beijing 100084, China

²UMich Joint Institute, Shanghai Jiao Tong University, Shanghai 200240, China

³Tandon School of Engineering, New York University, Brooklyn, NY 11201, USA

⁴Department of Industrial Engineering and Operations Research, University of California, Berkeley, Berkeley, CA 94720, USA

⁵Department of Civil and Environmental Engineering, Northeastern University, Boston, MA 02115, USA

⁶Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Corresponding author: Li Jin, li.jin@sjtu.edu.cn.

Abstract. Short service suspensions are common in public transit systems, but their operational impacts remain difficult to quantify. We develop an analytical framework for measuring the resilience of a transit line under short random service suspensions. Vehicle movement is represented by a two state process in which vehicles either travel normally or stop during a suspension, and the induced stochastic headways enter a bulk service queueing model with finite vehicle capacity and passenger carryover. The model yields two classes of resilience indicators. Stability conditions determine whether station queues remain bounded, while closed form expressions characterize the mean and variance of station level queue length and waiting time. We construct an independent renewal approximation for headways whose common marginal distribution is obtained by taking the positive part of a raw headway formed from the incident adjusted scheduled headway and the difference between two independent compound Poisson exponential variables. The renewal approximation preserves the marginal effects of short suspensions while omitting serial dependence and delay propagation across multiple vehicles. Combining the resulting passenger arrival distribution with a Markov representation of passenger loads across stations allows the resilience indicators to be computed sequentially along the route. The stability condition identifies how incident frequency, incident duration, passenger demand, scheduled headway, vehicle capacity, and downstream available capacity jointly determine whether a station becomes unstable. Numerical experiments show that short suspensions disproportionately affect congested stations and that changes in incident duration and scheduled headway can dominate comparable changes in vehicle capacity. A recursive first in, first out simulation assesses the analytical approximations and clarifies the role of headway variability. The framework provides a tractable way to assess capacity, scheduled headway, and incident mitigation decisions for improving the resilience of transit lines exposed to frequent short disruptions.

Key words: public transit resilience; service suspension; headway variability; bulk service queueing; stability

1. Introduction

Public transit systems (PTSs) connect people with jobs, homes, and other daily activities and therefore play an important role in cities worldwide. Their service, however, is susceptible to

frequent unplanned delays and disruptions. For example, [Mo et al. \(2022\)](#) report an average of 75 incidents per day in the Chicago urban rail system, of which more than 75% last less than 5 minutes. These short suspensions may result from signal failures, passenger behavior, or infrastructure problems. Quantifying these effects is therefore important for assessing transit resilience.

In this study, we use two operational resilience metrics for PTSs, **queue length** and **waiting time**. Specifically, we model a PTS as a bulk service queue and aim to derive closed form expressions for the mean and variance of passenger queue length and waiting time at a station under random service suspensions. To this end, we derive a stability criterion for passenger queues under a renewal representation of suspension induced headways. The representation captures throughput loss when a nonpositive raw headway is converted to bunching and the lost separation is not recovered in a later headway. We also characterize the steady state distribution of passenger queues, which naturally leads to the quantification of these metrics.

Queueing behavior at a public transit station is commonly represented by a bulk service queue, in which a vehicle serves waiting passengers up to its available capacity ([Powell 1981](#)). Most transit applications analyze one station, while [Islam et al. \(2014\)](#) extended the analysis to a route through a Markov representation of downstream passenger flows. These studies focus on normal operations. Models of bulk service under disruptions instead represent a server with operating and repair states and usually assume a fixed batch size ([Madan 1989](#), [Krishnamoorthy, Pramod, and Chakravarthy 2014](#)). That assumption does not represent a transit route where the available boarding capacity varies with the onboard load and passenger alighting. Mapping a server breakdown to route operations is also difficult because passenger flows link the service capacities of consecutive stations along the route.

To fill these research gaps, we develop a bulk service queueing framework that links suspension induced headways, stochastic passenger arrivals, station queues, and the vehicle capacity propagated from upstream stations. The route calculation allows available boarding capacity to vary with onboard load and passenger alighting. Tracking these interactions station by station shows how short disruptions erode transit resilience through queue growth, waiting time variability, and station instability. These measures can inform future control and planning strategies for public transit.

Within this framework, we explicitly model random service suspensions in a single route PTS. Such a system represents a bus route or one direction of a rail line and is a basic element of more complex public transit networks. Each vehicle may experience random suspensions while traveling along the route. Section 3.2 explains the operational meaning of this assumption, shows how it

corresponds to many real world situations, and presents the assumption as a first step toward a general incident representation for PTSs.

Our contributions are summarized as follows.

1. *We develop an analytical bulk service queueing framework for quantifying line level public transit resilience under short random service suspensions.* In particular, we represent service suspensions through a vehicle speed process with normal and disruption states. The speed process gives a direct operational meaning to the abstract notion of a server breakdown in public transit operations under short suspensions.

2. *We develop an analytically tractable renewal approximation for headways under random service suspensions.* The common marginal distribution is constructed as the positive part of a raw headway formed from the incident adjusted scheduled headway and the difference between two independent compound Poisson exponential variables. Under the approximation, headways are independent across vehicle runs, and departure epochs accumulate from these nonnegative headways. As a result, the analytical service sequence prevents overtaking without imposing a pathwise first in, first out (FIFO) recursion. We then use a censored normal approximation with a point mass at zero to obtain a closed form moment generating function and derive the probability generating function (PGF) and moments of passenger arrivals within a headway. This is a new analytical contribution to bulk service queueing under short service suspensions.

3. *We develop a route level probability propagation model for inter station passenger flows.* The model recursively links passenger alighting, available vehicle capacity, station queues, and departing vehicle loads. These links allow the resilience indicators to be computed sequentially at every station along the route.

4. *We develop an interpolation based root solving algorithm that makes the analytical queueing indicators computationally implementable for realistic vehicle capacities.* The algorithm finds all complex roots required by the transform based bulk service queue model, addressing a numerical bottleneck that standard single start nonlinear solvers and existing root search heuristics may miss.

The rest of this paper is organized as follows. Section 2 reviews the literature on the bulk service queue problem, random service disruptions, and queueing models for PTSs. Section 3 presents the model settings for a single route system with random service suspensions. Section 4 presents the major analytical results and their derivations. Section 5 provides numerical examples to illustrate the theoretical results and assesses the proposed approach using simulation. Section 6 concludes the paper and discusses future research directions.

2. Literature review

2.1. Bulk service queues

Bulk service queues serve customers in groups rather than individually. Bailey (1954) formulated an early fixed capacity model with Poisson arrivals and an embedded Markov chain at service completion epochs. For transportation terminals, Powell (1981) developed transform methods for random passenger arrivals and bulk service. These models provide the queueing foundation for capacity constrained passenger boarding.

2.2. Random service disruptions

Models of queues with service breakdowns typically alternate between operating and repair states. Madan (1989) introduced this structure to a fixed batch bulk service queue, while Krishnamoorthy, Pramod, and Chakravarthy (2014) review the broader breakdown literature. These models are not directly applicable to a transit route because the service capacity at a station depends on the onboard load propagated from upstream stations.

2.3. Queueing models in public transit systems

Transit queueing analyses are usually conducted at the station level. Under regular service, suppose that passengers arrive according to a Poisson process, can board the first arriving vehicle, and face a deterministic headway H . Their mean waiting time is then

$$\mathbb{E}[W] = H/2, \quad (1)$$

where H is the service headway and W is the passenger waiting time. With stochastic headways, Osuna and Newell (1972) obtained

$$\mathbb{E}[W] = \frac{1}{2} \cdot \left[\mathbb{E}[H] + \frac{\text{Var}[H]}{\mathbb{E}[H]} \right], \quad (2)$$

where $\mathbb{E}[H]$ and $\text{Var}[H]$ are the expectation and variance of headways, respectively. When headways are deterministic, $\text{Var}[H] = 0$, and the model reduces to Eq. 1.

Both Eqs. 1 and 2 assume that every passenger can board the first vehicle. Once capacity becomes binding, however, some passengers are left behind and must wait for a later vehicle (Mo et al. 2020). Transform based bulk service models account for this additional delay and provide closed form moments of queue length and waiting time under capacity constraints (Powell 1985).

Extending these models from one station to a route requires the downstream propagation of passenger loads and available capacity. Hickman (2001) characterized route level headway and

passenger flow dynamics, and Islam et al. (2014) combined this structure with a bulk service model. Their formulation assumes Erlang headways and does not jointly propagate headway correlation and vehicle capacity. Empirical headway distributions can differ substantially from the Erlang distribution (Bellei and Gkoumas 2010). Our model instead derives a suspension induced headway distribution and links it to downstream capacity propagation.

2.4. Service interruptions in public transit systems

Research on public transit interruptions includes impact analysis and operations control. Impact studies measure observed operational effects (Mo et al. 2022). Control studies examine responses such as substitute bus services (Jin, Teo, and Odoni 2016) and passenger path recommendations under demand or behavior uncertainty (Mo et al. 2023, 2025). The present study complements these approaches by deriving analytical resilience indicators, including stability conditions and station level moments of queue length and waiting time, within a route level capacity constrained model.

3. Model

This section defines the transit line, passenger flow primitives, and stochastic suspension process used in the analysis. We begin with the single route bulk service setting and describe how vehicles and passengers interact at each station. We then represent short service suspensions with a two state vehicle speed process and derive the stochastic headway used in the station level queueing analysis.

3.1. Single route public transit system and vehicle movements

Consider the single route PTS with N stations shown in Figure 1. Vehicles depart from a transportation hub, denoted by station 0, and visit stations 1 through N in sequence. At station n , passenger arrivals follow a Poisson process with a fixed rate $\lambda^{(n)}$ during the period of interest. When a vehicle arrives, each onboard passenger independently alights with probability $\alpha^{(n)}$. Consequently, the number of alighting passengers at station n follows a binomial distribution. Poisson arrivals and binomial alighting are common assumptions in the public transit literature (Hickman 2001). Because the paper focuses on short service suspensions, we assume that passengers continue to wait and do not renege. The assumption is consistent with evidence that passenger abandonment becomes more important during substantially longer delays (Rahimi et al. 2019). Models of longer disruptions could relax the assumption by incorporating balking and reneging. Section EC.1 of the Electronic Companion (EC) appendix summarizes the modeling assumptions and their operational interpretations for the current setting.

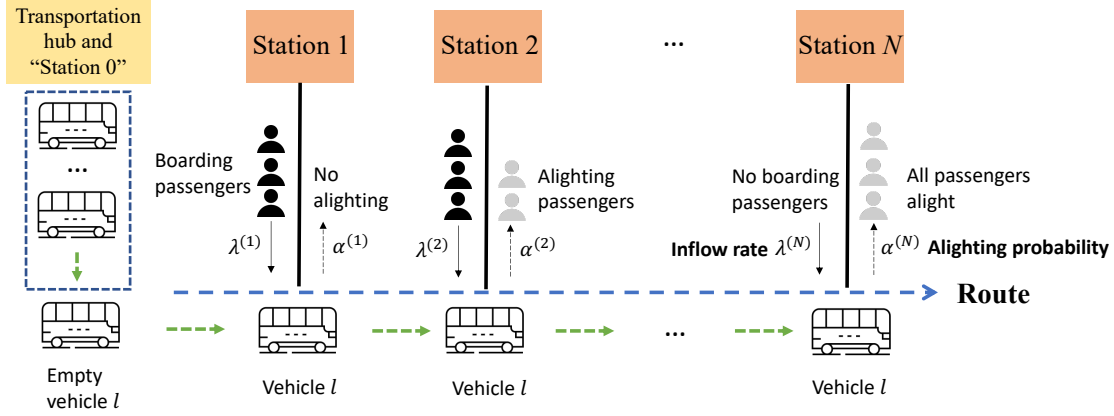


Figure 1 Schematic presentation of a single route public transit system

Let $l = 1, 2, \dots$ denote the vehicle run index, with a smaller value indicating an earlier dispatch. Figure 2 summarizes the vehicle and passenger interactions at station n over time. Let $t_A^{(n,l)}$ be the arrival time of vehicle l at station n and $t_D^{(n,l)}$ its departure time. For $l = 2, 3, \dots$, we use $H^{(n,l)}$ provisionally for the service interval between vehicle runs $l - 1$ and l . Section 3.3 distinguishes its stationwise FIFO realization from the analytical renewal approximation. When a vehicle arrives, onboard passengers alight before queueing passengers board in their order of arrival. Let $Q^{(n,l)}$ be the number of queueing passengers when vehicle l arrives at station n , $R^{(n,l)}$ the number left behind when it departs, and $Y^{(n,l)}$ the number arriving between $t_D^{(n,l)}$ and $t_A^{(n,l+1)}$. By definition,

$$Q^{(n,l+1)} = R^{(n,l)} + Y^{(n,l)}. \quad (3)$$

Following Powell (1981), we assume that dwell time is negligible relative to travel time and that no passenger arrives during dwell. Under this assumption, the number of passenger arrivals depends on the time between consecutive vehicles. For any service interval of length $h \geq 0$, the conditional arrival count satisfies

$$Y^{(n,l)} \mid (\text{service interval} = h) \sim \text{Poi}(\lambda^{(n)} h). \quad (4)$$

Thus, $Y^{(n,l)}$ counts passenger arrivals during the realized service interval.

To describe the same interaction from the vehicle side, let $S^{(n,l)}$ be the number of available spaces after passengers alight from vehicle l at station n , and let $G^{(n,l)}$ be the number of passengers who remain onboard. By definition,

$$G^{(n,l)} := C - S^{(n,l)}, \quad (5)$$

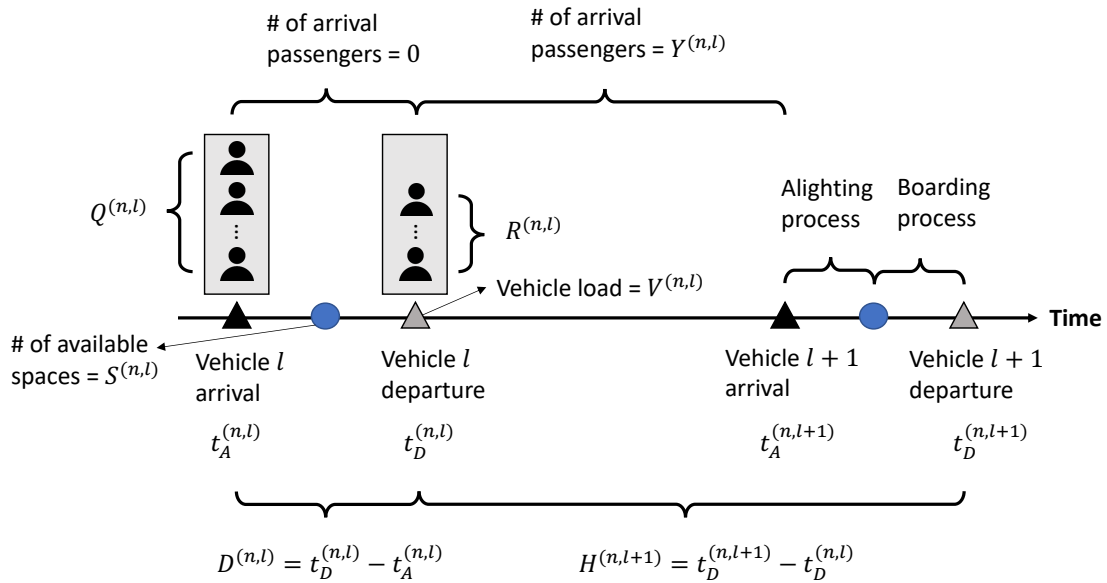


Figure 2 Vehicle and passenger interactions at station n over time

where C is the vehicle capacity. Let $V^{(n,l)}$ denote the vehicle load when vehicle l departs station n , which is also its load when it arrives at station $n + 1$. Conditional on $V^{(n-1,l)}$, the number of passengers who alight from vehicle l at station n follows a binomial distribution.

$$\left(V^{(n-1,l)} - G^{(n,l)} \right) \mid V^{(n-1,l)} \sim \text{Bin}(V^{(n-1,l)}, \alpha^{(n)}). \quad (6)$$

3.2. Random service suspensions and vehicle speed profile

We assume that random service suspensions may occur as a vehicle travels through the system. The red curve in Figure 3a shows a possible speed profile for vehicle l between stations n and $n + 1$ under such disturbances. Each incident slows or stops the vehicle. In bus systems, incidents may result from traffic congestion, accidents, vehicle engine problems, or driver and passenger behavior. In rail systems, they may result from signal or infrastructure failures and driver or passenger behavior. The speed profile provides a general representation of incidents, interruptions, suspensions, or disruptions that impede vehicle movement. We refer to the triggering event as an incident and to the resulting interruption of vehicle movement as a service suspension.

An actual vehicle speed profile under service suspensions can be complicated. For analytical tractability, we approximate it by the impulse function shown by the blue line in Figure 3a. The impulse approximation separates the trajectory into traveling and stopping phases, which we call the normal and disruption states. In the normal state, a vehicle travels at a constant speed. When an incident occurs, the vehicle stops immediately and enters the disruption state. Over a sufficiently

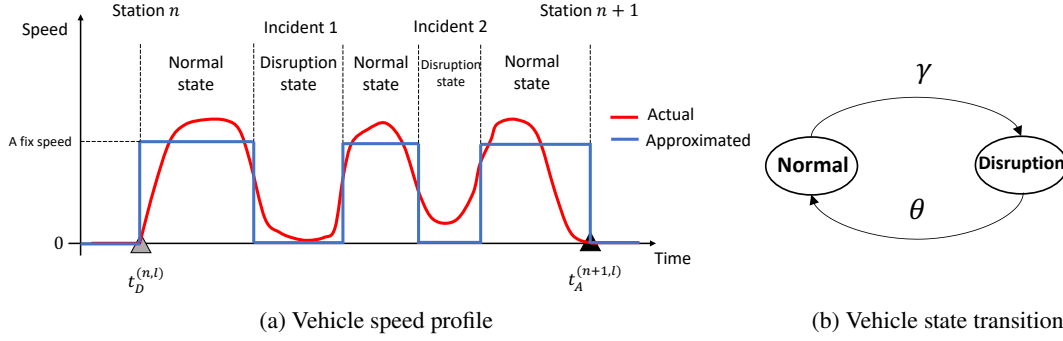


Figure 3 Representation of random service suspensions

small interval Δ , the probability of an incident is $\gamma\Delta$. Once the vehicle enters the disruption state, the incident duration follows an exponential distribution with rate θ and mean $1/\theta$. The vehicle state therefore follows the two state Markov process in Figure 3b with state space $\{\text{Normal}, \text{Disruption}\}$. Equivalently, the durations of the normal and disruption states are exponentially distributed with transition rates γ for incident occurrence and θ for service recovery.

The impulse function provides a first step toward a more general incident representation. Different incident categories could be represented by normal and disruption states with heterogeneous occurrence rates and durations, producing a richer speed profile.

3.3. Headway under random service suspensions

Under the assumed impulse function speed profile, all vehicles have the same fixed travel speed in the normal state. Consequently, when no incident occurs, all stations have the same deterministic headway $\bar{H}$, which we call the nominal scheduled headway. To relate this headway to route planning, let $\bar{E}$ denote the route cycle time required for a vehicle to travel from the transportation hub to the last station and return to the hub. Given route fleet size $\bar{F}$, the scheduled headway is $\bar{H} := \bar{E}/\bar{F}$. Random service suspensions increase the expected route cycle time and therefore require a planning response. A transit agency can maintain the scheduled headway $\bar{H}$ by increasing the route fleet size, or it can maintain the same fleet size by increasing the scheduled headway. We consider the second response because it makes the effects of incidents on headway and service performance explicit, which is the focus of this paper.

Following this planning response, we assume that the transit agency has an estimate $\bar{D}$ of the average incident delay in the cycle time. Let $I^{(n,l)}$ be the total duration of all incidents while vehicle l travels from the transportation hub to station n . Then $\mathbb{E}[I^{(N,l)}]$ is the expected incident duration

for travel from the hub to the last station N . Assuming the same road conditions in both directions, the total estimated delay for the cycle trip is

$$\overline{D} := 2 \cdot \mathbb{E}[I^{(N,l)}]. \quad (7)$$

Some transit agencies may plan for a delay above the mean, such as the 85th percentile. In that case, the cycle delay could be formulated as a more general function of the incident duration distribution. We use the expected delay in Eq. 7 to keep the subsequent analysis transparent. After incorporating this delay, the incident adjusted scheduled headway is $\overline{H}^{\text{Adj}} := (\overline{E} + \overline{D})/\overline{F} = \overline{H} + 2\mathbb{E}[I^{(N,l)}]/\overline{F}$. The term $2\mathbb{E}[I^{(N,l)}]/\overline{F}$ is therefore the scheduled headway adjustment associated with incidents. Because $I^{(N,l)}$ is identically distributed across vehicle runs, the adjusted scheduled headway does not depend on the vehicle run index. The single route PTS then dispatches vehicles on time at intervals of $\overline{H}^{\text{Adj}}$.

Let $T^{(n)}$ be the travel time from the transportation hub to station n when there is no incident. This travel time is fixed under the constant speed assumption. Since dwell time is negligible, the arrival and departure times at a station are identical in the model. Taking the dispatch time of vehicle 1 as time zero, the free departure time of vehicle l at station n is

$$\hat{t}_D^{(n,l)} := (l-1)\overline{H}^{\text{Adj}} + T^{(n)} + I^{(n,l)}. \quad (8)$$

The free departure time in Eq. 8 is the time that would result if vehicle order were unrestricted. Under this construction, a sufficiently long incident affecting vehicle $l-1$ can yield $\hat{t}_D^{(n,l)} < \hat{t}_D^{(n,l-1)}$. Vehicle l would then appear to overtake vehicle $l-1$, although vehicle $l-1$ was dispatched first.

To rule out such reversals, we assume that vehicles serve each station in dispatch order and that overtaking is not allowed. The FIFO rule represents rail lines where passing is infeasible and bus services where passing is prohibited or uncommon at the route scale. It also keeps the vehicle index aligned with the service order observed by waiting passengers. If the free departure time of vehicle l precedes the realized departure time of vehicle $l-1$, vehicle l is held until vehicle $l-1$ departs. The realized departure times at station n therefore satisfy

$$t_D^{(n,1)} := \hat{t}_D^{(n,1)}, \quad t_D^{(n,l)} := \max \left\{ \hat{t}_D^{(n,l)}, t_D^{(n,l-1)} \right\}, \quad \forall l \geq 2. \quad (9)$$

Under Eq. 9, the stationwise FIFO headway is $H_{\text{FIFO}}^{(n,l)} := t_D^{(n,l)} - t_D^{(n,l-1)} = \max\{0, \hat{t}_D^{(n,l)} - t_D^{(n,l-1)}\}$. This headway depends on the realized departure time of the preceding vehicle and can therefore

inherit delays from several earlier vehicles at the same station. Direct use of this recursion would therefore require the departure times of multiple vehicle runs to be analyzed jointly.

To avoid this recursive dependence, we impose the following approximation before constructing the analytical headway distribution.

ASSUMPTION 1 (Independent renewal headways). *At each station n , the analytical headways $\{H^{(n,l)} : l = 2, 3, \dots\}$ are identically and independently distributed. Each headway is independent of all prior headways and of the queue and vehicle load states observed before the interval begins. The incident draws used to construct a headway are also independent of the exogenous passenger arrival process. The common marginal distribution is constructed from the free departure times of a generic pair of consecutive vehicles, with two new independent incident durations drawn for every headway in the renewal sequence.*

Although the incident draws are independent of the exogenous Poisson passenger arrival process, the number of passengers arriving within an interval depends on its length through Eq. 4. Separately, on time dispatch supports the assumed independence of the analytical headway from prior queue and vehicle load states. Independence across vehicle runs is an analytical approximation that is most defensible when the total incident delay of one vehicle rarely extends across several scheduled headways. The condition is consistent with service suspensions that are short relative to the scheduled headway. The approximation is also suitable for headway based operations in which holding or dispatch control resets the next service interval from the current departure (Daganzo 2009). It becomes less suitable when a long delay propagates through many subsequent vehicles. Section 5 evaluates this limitation using a simulation that retains the recursive FIFO dynamics in Eq. 9.

Under Assumption 1, the common marginal distribution is obtained from the difference between the free departure times of a generic pair of consecutive vehicles. Accordingly, subtracting the two free departure times in Eq. 8 gives the raw headway

$$\widehat{H}^{(n,l)} := \hat{t}_D^{(n,l)} - \hat{t}_D^{(n,l-1)} = \overline{H}^{\text{Adj}} + I^{(n,l)} - I^{(n,l-1)}, \quad l = 2, 3, \dots \quad (10)$$

The deterministic term in Eq. 10 is $\overline{H}^{\text{Adj}} = \overline{H} + 2\mathbb{E}[I^{(N,l)}]/\overline{F}$. For one marginal headway calculation, the two incident durations are independent. In a physical vehicle sequence, adjacent raw headways share one incident duration and are therefore dependent. The renewal model draws a new pair of incident durations for each analytical headway. If the difference between the two durations is

sufficiently negative, $\widehat{H}^{(n,l)}$ becomes negative and cannot represent elapsed time between service events. We therefore define the analytical headway as

$$H^{(n,l)} := \max\{0, \widehat{H}^{(n,l)}\}, \quad l = 2, 3, \dots \quad (11)$$

Equation 11 is not algebraically equivalent to the recursive FIFO rule in Eq. 9. A nonpositive raw headway becomes zero and represents bunching, but the next headway is redrawn instead of inheriting the accumulated delay. The approximation preserves vehicle order through nonnegative analytical service intervals, but it omits serial dependence and the propagation of accumulated delay to later vehicles.

4. Analysis

The analysis derives the stability condition in Proposition 12 and the mean and variance of passenger queue length and waiting time in Propositions 5 and 6. Figure 4 places these station level results within the route calculation by summarizing the propagation of the distributions of $S^{(n,l)}$, $V^{(n,l)}$, and $Q^{(n,l)}$. The calculation consists of three linked updates.

- Given the distribution of $V^{(n-1,l)}$, calculate the distribution of the available spaces $S^{(n,l)}$ after alighting at station n , as shown in Section 4.1.
- Given the distribution of $S^{(n,l)}$, calculate the distribution of the queue $Q^{(n,l)}$ and the moments of queue length and waiting time, as shown in Section 4.3.
- Given the distributions of $S^{(n,l)}$ and $Q^{(n,l)}$, calculate the distribution of the departing vehicle load $V^{(n,l)}$, as shown in Section 4.2.

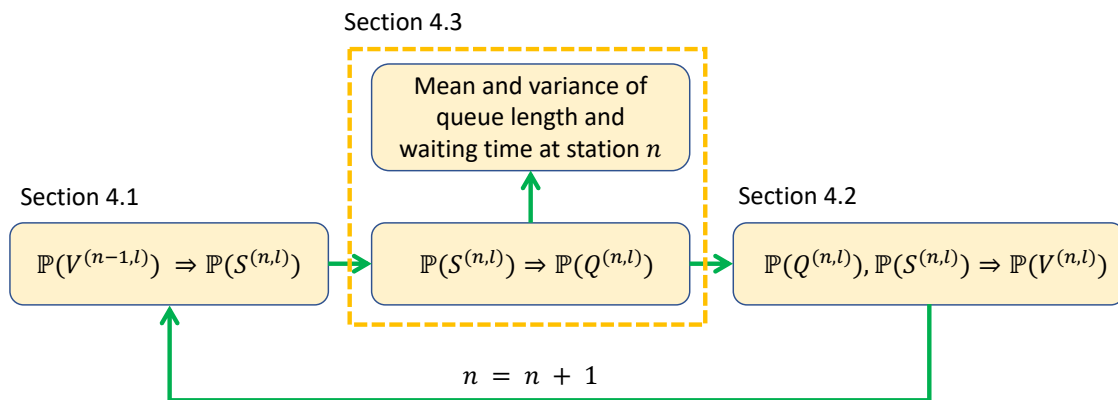


Figure 4 Analysis framework

Together, these components yield the distributions of $S^{(n,l)}$, $Q^{(n,l)}$, and $V^{(n,l)}$ for all $n = 1, \dots, N$. The recursion starts from $V^{(0,l)} = 0$, which represents an empty vehicle departing the hub for the first station. The analysis focuses on steady state distributions as $l \rightarrow \infty$.

4.1. Available vehicle space distribution

This subsection derives the steady state distribution of $S^{(n,l)}$ from that of $V^{(n-1,l)}$. Define $v_k^{(n,l)} := \mathbb{P}(V^{(n,l)} = k)$, $s_k^{(n,l)} := \mathbb{P}(S^{(n,l)} = k)$, and $g_k^{(n,l)} := \mathbb{P}(G^{(n,l)} = k)$ for $k = 0, 1, \dots, C$. When the steady state probabilities exist, define $v_k^{(n)} := \lim_{l \rightarrow \infty} v_k^{(n,l)}$, $s_k^{(n)} := \lim_{l \rightarrow \infty} s_k^{(n,l)}$, and $g_k^{(n)} := \lim_{l \rightarrow \infty} g_k^{(n,l)}$. Let $V^{(n)}$, $S^{(n)}$, and $G^{(n)}$ denote random variables with these steady state distributions. Section 4.4 establishes the stability condition under which the distributions exist.

PROPOSITION 1 (Distribution of available vehicle space $S^{(n)}$). *For each $n = 1, \dots, N$, let $v^{(n-1)} := [v_0^{(n-1)}, \dots, v_C^{(n-1)}] \in \mathbb{R}^{C+1}$ and $s^{(n)} := [s_0^{(n)}, \dots, s_C^{(n)}] \in \mathbb{R}^{C+1}$. Let $g^{(n)} := [g_0^{(n)}, \dots, g_C^{(n)}] \in \mathbb{R}^{C+1}$ denote the distribution of the vehicle load after alighting, with $g^{(n)} = v^{(n-1)} A^{(n)}$. The available space distribution is then given by the following reversal of the vehicle load index.*

$$s_k^{(n)} = g_{C-k}^{(n)} \quad \forall k = 0, 1, \dots, C. \quad (12)$$

The matrix $A^{(n)}$ has dimension $(C+1) \times (C+1)$. Its element $a_{ij}^{(n)}$ in row i and column j is the probability that j passengers remain on board after alighting, conditional on i passengers being on board before alighting. It is defined as

$$a_{ij}^{(n)} := \begin{cases} \binom{i}{j} (1 - \alpha^{(n)})^j (\alpha^{(n)})^{i-j}, & \text{if } 0 \leq j \leq i \leq C, \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

4.2. Vehicle load distribution

This subsection states the steady state update for $V^{(n,l)}$ in terms of the distributions of $G^{(n,l)}$ and $Q^{(n,l)}$; Section 4.3 subsequently derives the required queue distribution. Combining these two marginal distributions requires the following approximation for route propagation.

ASSUMPTION 2 (Marginal route decoupling). *At each station n , the steady state remaining onboard load $G^{(n)}$ and the steady state local passenger queue $Q^{(n)}$ are independent. Therefore, $\mathbb{P}(G^{(n)} = i, Q^{(n)} = k) = \mathbb{P}(G^{(n)} = i) \mathbb{P}(Q^{(n)} = k)$ for all feasible i and k .*

Assumption 2 is exact at the first station because vehicles arrive from the hub without passengers and $G^{(1)} = 0$. At a downstream station, the remaining onboard load results from boarding at upstream stations followed by binomial alighting, whereas the local queue results from the station specific Poisson arrival process and passengers previously left behind at that station. These quantities therefore arise from different passenger flow components. Independent renewal headways, alighting at intermediate stations, and repeated queue clearance under stable operation weaken the dependence

induced by common headways and upstream capacity constraints. The approximation is most appropriate for short suspensions and stable stations, and it becomes weaker when long disruptions or persistent oversaturation jointly produce high onboard loads and long local queues. It permits the two marginal distributions to be propagated separately along the route. Section 5 assesses the approximation through recursive simulation.

Define $q_k^{(n,l)} := \mathbb{P}(Q^{(n,l)} = k)$ and, when the steady state distribution exists, $q_k^{(n)} := \lim_{l \rightarrow \infty} q_k^{(n,l)}$. Let $Q^{(n)}$ denote a random variable with probabilities $\mathbb{P}(Q^{(n)} = k) = q_k^{(n)}$. The vector $q_{0:C-1}^{(n)} := [q_0^{(n)}, \dots, q_{C-1}^{(n)}] \in \mathbb{R}^C$ contains its first C probabilities.

PROPOSITION 2 (Distribution of vehicle load $V^{(n)}$). *Under Assumption 2, the distribution of $V^{(n)}$ at each station $n = 1, \dots, N$, given $g^{(n)}$ and $q_{0:C-1}^{(n)}$, satisfies*

$$v^{(n)} = g^{(n)} B^{(n)}, \quad (14)$$

where $b_{ij}^{(n)}$, the element in row i and column j of $B^{(n)}$, is defined as follows.

$$b_{ij}^{(n)} := \begin{cases} q_{j-i}^{(n)}, & \text{if } 0 \leq i \leq j < C \\ 1 - \sum_{k=0}^{C-i-1} q_k^{(n)}, & \text{if } j = C \text{ and } 0 \leq i < C \\ 1, & \text{if } i = j = C \\ 0, & \text{otherwise} \end{cases} \quad i, j = 0, 1, \dots, C. \quad (15)$$

4.3. Queueing analysis at a station

Given the distribution $s^{(n)} := [s_0^{(n)}, \dots, s_C^{(n)}] \in \mathbb{R}^{C+1}$ of $S^{(n)}$, this subsection identifies the characteristic roots and determines $q_{0:C-1}^{(n)}$ and the mean and variance of passenger queue length and waiting time. The transform calculations in this subsection are conditional on station stability, so the steady state queue and its PGF are well defined. Section 4.4 establishes the stability condition independently from the transient queue recursion. For compact notation, let $Y^{(n)}$ denote the steady state number of passenger arrivals within a headway at station n , and write $\bar{S}^{(n)} = \mathbb{E}[S^{(n)}]$. The quantities $\bar{\bar{S}}^{(n)}$ and $\bar{\bar{\bar{S}}}^{(n)}$ denote the second and third central moments of $S^{(n)}$, respectively, and the same convention applies to $Y^{(n)}$.

ASSUMPTION 3 (Exogenous station service sequence). *For each station n , once the marginal distribution $s^{(n)}$ has been propagated from upstream, the station level analytical model represents $\{S^{(n,l)} : l = 1, 2, \dots\}$ as independent draws from $s^{(n)}$. This sequence is independent of the analytical headways, local passenger arrivals, and local queue history. The difference $Y^{(n,l)} - S^{(n,l+1)}$ has a finite mean and is nonconstant when $\bar{Y}^{(n)} = \bar{S}^{(n)}$.*

4.3.1. Probability generating function and characteristic roots. We first derive the PGF of the steady state queue $Q^{(n)}$ and then identify the characteristic roots required by the queue probability calculation in Section 4.3.2. Throughout this station level derivation, $Q(z)$, $Y(z)$, $\text{NUM}(z)$, and $\text{DEN}(z)$ refer to station n , whose superscript is suppressed.

PROPOSITION 3 (PGF of $Q^{(n)}$). *Under Assumption 3, for each $n = 1, \dots, N$, the available space distribution $s^{(n)}$ yields the following PGF for the station queue $Q^{(n)}$.*

$$Q(z) = \frac{\sum_{u=0}^C s_u^{(n)} \left[\sum_{i=0}^u q_i^{(n)} (z^C - z^{C-u+i}) \right]}{\frac{z^C}{Y(z)} - \sum_{u=0}^C s_u^{(n)} z^{C-u}}, \quad (16)$$

where $Y(z)$ is the PGF of $Y^{(n)}$.

Let $\text{NUM}(z)$ and $\text{DEN}(z)$ denote the numerator and denominator of $Q(z)$ in Eq. 16. Because $Q(z)$ is a PGF, it is analytic for $|z| < 1$ and bounded for $|z| \leq 1$. Every nonunit root of $\text{DEN}(z)$ in the closed unit disk must therefore be canceled by a root of $\text{NUM}(z)$ with the same multiplicity. The unit root is also a root of $\text{NUM}(z)$ because direct substitution gives $\text{NUM}(1) = 0$. Since $\text{NUM}(z)$ is a polynomial of degree at most C , the C characteristic roots counted according to multiplicity determine its factorization up to a scale factor. Section 4.3.2 uses this factorization to solve for $q_0^{(n)}, \dots, q_{C-1}^{(n)}$.

To locate the required roots, we multiply $\text{DEN}(z)$ by $Y(z)$ to obtain the characteristic function

$$\mathcal{F}^{(n)}(z) := z^C - Y(z) \sum_{u=0}^C s_u^{(n)} z^{C-u}. \quad (17)$$

The result of [Adan, van Leeuwen, and Winands \(2006\)](#) implies that $\mathcal{F}^{(n)}(z)$ has exactly C roots in the closed unit disk when roots are counted according to multiplicity. Because $s_C^{(n)} > 0$ and $Y(0) > 0$, zero is not a root of $\mathcal{F}^{(n)}(z)$. At every nonzero root, $Y(z)$ is nonzero, so the same root and multiplicity apply to $\text{DEN}(z)$. In particular, $z = 1$ is a simple root because $[\mathcal{F}^{(n)}]'(1) = \bar{S}^{(n)} - \bar{Y}^{(n)} > 0$. Interior roots need not be distinct ([Oblakova et al. 2019](#)). We therefore solve $\mathcal{F}^{(n)}(z) = 0$ for the distinct root locations $\zeta_1, \dots, \zeta_J$ in the closed unit disk, where J denotes the number of distinct locations. Let $m_1, \dots, m_J$ denote their corresponding multiplicities, which satisfy $\sum_{j=1}^J m_j = C$. For use in the queue calculation, we list the roots as $z_0^*, \dots, z_{C-1}^*$ according to multiplicity and adopt the convention $z_0^* = 1$. Section 4.3.2 uses this list to determine $q_0^{(n)}, \dots, q_{C-1}^{(n)}$, while Section 4.3.6 presents the numerical procedure for constructing it.

4.3.2. Queue length distribution. Section 4.3.1 identifies the C characteristic roots required to factor the numerator of $Q(z)$. We now translate that factorization into the queue probabilities $q_0^{(n)}, \dots, q_{C-1}^{(n)}$. First, the normalization condition $Q(1) = 1$ determines the scale factor. Matching the resulting polynomial coefficients then yields the required probabilities.

PROPOSITION 4 (Distribution of queue length $Q^{(n)}$). *For every station $n = 1, \dots, N$ with $s_C^{(n)} > 0$, the distribution of $S^{(n)}$, the roots $z_0^*, \dots, z_{C-1}^*$ listed according to multiplicity, and $\bar{Y}^{(n)}$ determine the probabilities $q_{0:C-1}^{(n)}$ through*

$$q_0^{(n)} = \frac{1}{s_C^{(n)}} (\bar{S}^{(n)} - \bar{Y}^{(n)}) \prod_{i=1}^{C-1} \frac{z_i^*}{z_i^* - 1}, \quad q_{0:C-1}^{(n)} = \tilde{\eta}^{(n)} (\Lambda^{(n)})^{-1}. \quad (18)$$

where the coefficients $\eta_j^{(n)}$ are defined by $\sum_{j=0}^C \eta_j^{(n)} z^j := \prod_{i=0}^{C-1} (1 - z/z_i^*)$. Let $\tilde{\eta}^{(n)} := [s_C^{(n)} q_0^{(n)} \eta_0^{(n)}, s_C^{(n)} q_0^{(n)} \eta_1^{(n)}, \dots, s_C^{(n)} q_0^{(n)} \eta_{C-1}^{(n)}] \in \mathbb{R}^C$ and

$$\Lambda^{(n)} := \begin{bmatrix} s_C^{(n)} & s_{C-1}^{(n)} & \cdots & s_2^{(n)} & s_1^{(n)} \\ 0 & s_C^{(n)} & \cdots & s_3^{(n)} & s_2^{(n)} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & s_C^{(n)} & s_{C-1}^{(n)} \\ 0 & \cdots & \cdots & 0 & s_C^{(n)} \end{bmatrix} \in \mathbb{R}^{C \times C}. \quad (19)$$

A root of multiplicity m appears m times in the defining product and in the root sums used below. Because the roots depend on station n , the coefficients $\eta_j^{(n)}$ carry the same station superscript.

The condition $s_C^{(n)} > 0$ holds at the first station because every vehicle arrives empty. At a downstream station, it holds whenever there is a positive probability that all onboard passengers alight or that the vehicle arrives empty. In a degenerate case with $s_C^{(n)} = 0$, the station level transform can instead use the largest u for which $s_u^{(n)} > 0$ as the effective maximum service capacity without changing the queue process.

4.3.3. Analytical formulation of mean and variance of queue length and waiting time. Once $q_0^{(n)}, \dots, q_{C-1}^{(n)}$ have been obtained, the PGF $Q(z)$ is fully determined. We can therefore recover the expectation and variance of queue length at station n from its first two derivatives as follows.

$$\mathbb{E}[Q^{(n)}] = \sum_{k=0}^{\infty} k q_k^{(n)} = \left. \frac{dQ(z)}{dz} \right|_{z=1}, \quad (20)$$

$$\text{Var}[Q^{(n)}] = \mathbb{E}[(Q^{(n)})^2] - \mathbb{E}[Q^{(n)}]^2 = \left. \frac{d^2 Q(z)}{dz^2} \right|_{z=1} + \mathbb{E}[Q^{(n)}] - \mathbb{E}[Q^{(n)}]^2. \quad (21)$$

PROPOSITION 5 (Mean and variance of queue length). *For each $n = 1, \dots, N$, given the distribution of $S^{(n)}$ and the PGF $Y(z)$, the mean and variance of $Q^{(n)}$ are*

$$\mathbb{E}[Q^{(n)}] = \frac{\bar{\bar{S}}^{(n)} + \bar{\bar{Y}}^{(n)} + (\bar{S}^{(n)} - \bar{Y}^{(n)})[1 + 2(\bar{S}^{(n)} - C)] - (\bar{S}^{(n)} - \bar{Y}^{(n)})^2}{2(\bar{S}^{(n)} - \bar{Y}^{(n)})} + \sum_{i=1}^{C-1} \frac{1}{1 - z_i^*}, \quad (22)$$

$$\begin{aligned} \text{Var}[Q^{(n)}] = & \frac{1}{12(\bar{S}^{(n)} - \bar{Y}^{(n)})^2} \left[-4(\bar{\bar{S}}^{(n)} - \bar{\bar{Y}}^{(n)})(\bar{S}^{(n)} - \bar{Y}^{(n)}) + 3(\bar{\bar{S}}^{(n)} + \bar{\bar{Y}}^{(n)})^2 \right. \\ & \left. - [6(\bar{\bar{S}}^{(n)} - \bar{\bar{Y}}^{(n)}) - 1](\bar{S}^{(n)} - \bar{Y}^{(n)})^2 - (\bar{S}^{(n)} - \bar{Y}^{(n)})^4 \right] - \sum_{i=1}^{C-1} \frac{z_i^*}{(1 - z_i^*)^2}. \end{aligned} \quad (23)$$

We next express waiting time in terms of the queue length observed at an arbitrary time.

PROPOSITION 6 (Mean and variance of waiting time). *For each $n = 1, \dots, N$ with $\lambda^{(n)} > 0$, given the distribution of $S^{(n)}$ and the PGF $Y(z)$, the mean and variance of waiting time $W^{(n)}$ are*

$$\mathbb{E}[W^{(n)}] = \frac{\bar{Q}_t^{(n)}}{\lambda^{(n)}}, \quad (24)$$

$$\text{Var}[W^{(n)}] = \frac{\bar{\bar{Q}}_t^{(n)} - \bar{Q}_t^{(n)}}{(\lambda^{(n)})^2}, \quad (25)$$

where $\bar{Q}_t^{(n)}$ is the queue length at an arbitrary time, whereas $Q^{(n)}$ is observed at a vehicle arrival. To make the distinction explicit, we write $\bar{Q}_t^{(n)} = \mathbb{E}[Q_t^{(n)}]$ and $\bar{\bar{Q}}_t^{(n)} = \text{Var}[Q_t^{(n)}]$. The following expressions link the arbitrary time moments to their arrival epoch counterparts.

$$\bar{Q}_t^{(n)} = \mathbb{E}[Q^{(n)}] - \bar{Y}^{(n)} + \frac{1}{2} \left(\bar{\bar{Y}}^{(n)} / \bar{Y}^{(n)} + \bar{Y}^{(n)} - 1 \right), \quad (26)$$

$$\bar{\bar{Q}}_t^{(n)} = \text{Var}[Q^{(n)}] - \bar{\bar{Y}}^{(n)} + \frac{1}{12(\bar{Y}^{(n)})^2} \left[4\bar{Y}^{(n)}\bar{\bar{\bar{Y}}}^{(n)} + 6(\bar{Y}^{(n)})^2\bar{\bar{\bar{Y}}}^{(n)} - (\bar{Y}^{(n)})^2 + (\bar{Y}^{(n)})^4 - 3(\bar{\bar{\bar{Y}}}^{(n)})^2 \right]. \quad (27)$$

Under the Poisson arrival and first come, first served assumptions, Eq. 24 applies Little's law, and Proposition 6 follows from Powell (1985).

REMARK 1. The queue length and waiting time expressions are consistent with the general $M/G^{[S]}/1$ bulk service results of Powell (1985). The contribution here is the specification of $Y(z)$ induced by random service suspensions and its integration with route level capacity propagation.

4.3.4. Headway distribution. As defined in Eq. 11, the analytical headway is $H^{(n,l)} := \max\{0, \widehat{H}^{(n,l)}\}$. For the steady state analysis, we suppress the vehicle run index and write the common marginal variables as $H^{(n)}$ and $\widehat{H}^{(n)}$. We derive the distribution of $H^{(n)}$ in two steps. We first characterize the distribution of the raw headway $\widehat{H}^{(n)}$ and then take its positive part to obtain the distribution of $H^{(n)}$. This distribution is required in Propositions 4 to 6 because Eq. 4 implies that $Y^{(n)} | H^{(n)}$ is Poisson with parameter $\lambda^{(n)} H^{(n)}$.

Equation 10 shows that the raw headway depends on the difference between two independent copies of the total incident duration for travel from the hub to station n . We therefore begin the derivation by characterizing the distribution of $I^{(n,l)}$.

PROPOSITION 7 (Distribution of incident duration). *The total incident duration $I^{(n,l)}$ for vehicle l traveling from the transportation hub to station n follows a compound Poisson exponential distribution with Poisson rate $\gamma T^{(n)}$ and exponential rate θ . It has the representation*

$$I^{(n,l)} := \sum_{i=1}^K X_i, \quad X_i \sim \text{Exp}(\theta), \quad i = 1, \dots, K, \quad K \sim \text{Poi}(\gamma T^{(n)}). \quad (28)$$

The moment generating function (MGF) of a compound Poisson exponential variable is

$$M_{I^{(n,l)}}(t) = \mathbb{E}[e^{tI^{(n,l)}}] = e^{\gamma T^{(n)}(\frac{\theta}{\theta-t}-1)}, \quad t < \theta. \quad (29)$$

Similarly, the MGF of $-I^{(n,l-1)}$ is

$$M_{-I^{(n,l-1)}}(t) = \mathbb{E}[e^{-tI^{(n,l-1)}}] = e^{\gamma T^{(n)}(\frac{\theta}{\theta+t}-1)}, \quad t > -\theta. \quad (30)$$

The MGF of $I^{(n,l)}$ gives $\mathbb{E}[I^{(n,l)}] = \gamma T^{(n)} / \theta$. Using this expectation in Eq. 10 and the independence of the two incident duration copies gives the raw headway MGF in Proposition 8.

PROPOSITION 8 (MGF of raw headway). *Under the setting of this study, for each $n = 1, \dots, N$, the MGF of $\widehat{H}^{(n)}$ can be expressed, for $|t| < \theta$, as*

$$M_{\widehat{H}^{(n)}}(t) = e^{t(\overline{H} + \frac{2\gamma T^{(N)}}{\theta \overline{F}})} e^{\gamma T^{(n)}(\frac{2t^2}{\theta^2 - t^2})}. \quad (31)$$

From the MGF of $\widehat{H}^{(n)}$, we obtain the mean and variance of the raw headway as

$$\mathbb{E}[\widehat{H}^{(n)}] = \overline{H} + \frac{2\gamma T^{(N)}}{\theta \overline{F}}, \quad (32)$$

$$\text{Var}[\widehat{H}^{(n)}] = \frac{4T^{(n)}\gamma}{\theta^2}. \quad (33)$$

REMARK 2. The results show that random suspensions can increase the mean and variance of the raw headway. The impact on its mean occurs through the cycle time adjustment at the planning stage. Its variance increases with a higher incident rate γ and a higher average incident duration $\frac{1}{\theta}$. The station specific marginal distributions also capture the downstream growth in headway variability described by [Hickman \(2001\)](#). In particular, $\text{Var}[\widehat{H}^{(n)}]$ increases with the station index n because $T^{(n)}$ increases.

Having characterized the raw headway, the second step applies the positive part transformation in Eq. 11. A direct MGF representation of the resulting headway involves nested integrals over the two compound Poisson distributions. Although this representation is exact, it does not reduce to a tractable finite expression for general parameter values. Moreover, the root calculations would require repeated numerical integration or infinite series evaluation at every complex argument. To avoid this computational burden, we first approximate the raw headway $\widehat{H}^{(n)}$ by a normal distribution with the same mean and variance and then take its positive part.

We use the normal approximation for two reasons. First, $I^{(n,l)}$ is a random sum of independent exponential durations. A normal approximation becomes more accurate as the expected incident count increases, consistent with the setting of frequent short disturbances. A related central limit approximation for headway disturbances appears in [Daganzo \(2009\)](#). Second, let $\widehat{H}_{\text{Normal}}^{(n)}$ denote a normal variable with the same mean and variance as $\widehat{H}^{(n)}$. Both variables are symmetric and have a third central moment of zero. Moreover, the MGF of $\widehat{H}_{\text{Normal}}^{(n)}$ is

$$M_{\widehat{H}_{\text{Normal}}^{(n)}}(t) = e^{t(\bar{H} + \frac{2\gamma T^{(N)}}{\theta F})} e^{\gamma T^{(n)} (\frac{2t^2}{\theta^2})}, \quad (34)$$

which has the same second order expansion around $t = 0$ as the exact MGF in Eq. 31. The fourth central moment is $48(T^{(n)}\gamma)^2/\theta^4$ for $\widehat{H}_{\text{Normal}}^{(n)}$ and $[48(T^{(n)}\gamma)^2 + 48T^{(n)}\gamma]/\theta^4$ for $\widehat{H}^{(n)}$. The exact raw headway may therefore have heavier tails. Figure 5 numerically compares the two distributions for several values of $T^{(n)}$, θ , and γ . The histograms are generated by sampling from the associated exponential and Poisson distributions. The normal distribution provides a close approximation in these examples while preserving the expected difference in tail behavior.

We next take the positive part of $\widehat{H}_{\text{Normal}}^{(n)}$. Denote the resulting headway approximation by $H_{\text{Normal}}^{(n)}$. It has support $[0, +\infty)$ and a probability mass at zero.

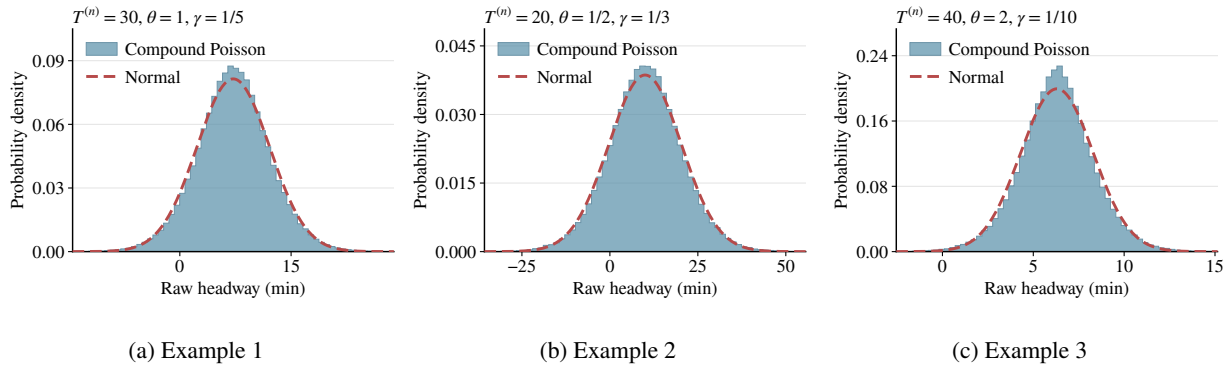


Figure 5 Numerical assessment of the normal approximation for the raw headway distribution

PROPOSITION 9 (MGF of approximated headway). For $\gamma > 0$ and $n = 1, \dots, N$, the MGF of $H_{Normal}^{(n)}$ is given in Eq. 35.

$$M_{H_{Normal}^{(n)}}(t) = \Phi\left(\frac{-(\bar{H}\theta + \frac{2\gamma T^{(N)}}{\bar{F}})}{2\sqrt{T^{(n)}\gamma}}\right) + e^{t(\bar{H} + \frac{2\gamma T^{(N)}}{\theta\bar{F}})} e^{\gamma T^{(n)}(\frac{2t^2}{\theta^2})} \left[1 - \Phi\left(\frac{-(\bar{H}\theta + \frac{2\gamma T^{(N)}}{\bar{F}})}{2\sqrt{T^{(n)}\gamma}} - \frac{2t\sqrt{T^{(n)}\gamma}}{\theta}\right)\right]. \quad (35)$$

Here $\Phi(\cdot)$ denotes the standard normal cumulative distribution function and its analytic continuation for complex arguments in the root calculation. When $\gamma = 0$, no incident duration shocks occur, so $H_{Normal}^{(n)} = \bar{H}$ almost surely and $M_{H_{Normal}^{(n)}}(t) = e^{t\bar{H}}$.

For compact notation, let $\mu := \bar{H} + 2\gamma T^{(N)} / (\theta\bar{F})$ and $\sigma := 2\sqrt{T^{(n)}\gamma} / \theta$ denote the mean and standard deviation of $\hat{H}_{Normal}^{(n)}$. Differentiating the MGF and using $1 - \Phi(-\mu/\sigma) = \Phi(\mu/\sigma)$ then gives the corresponding mean and variance of the approximated headway.

$$\mathbb{E}[H_{Normal}^{(n)}] = \mu \cdot \Phi\left(\frac{\mu}{\sigma}\right) + \sigma \cdot \phi\left(\frac{-\mu}{\sigma}\right), \quad (36)$$

$$\text{Var}[H_{Normal}^{(n)}] = \mu\sigma\phi\left(\frac{-\mu}{\sigma}\right) + \Phi\left(\frac{\mu}{\sigma}\right)(\mu^2 + \sigma^2) - \left(\mu\Phi\left(\frac{\mu}{\sigma}\right) + \phi\left(\frac{-\mu}{\sigma}\right)\sigma\right)^2, \quad (37)$$

where $\phi(\cdot)$ is the probability density function of a standard normal distribution. For $\gamma = 0$, these expressions are interpreted by continuity, giving mean $\bar{H}$ and variance zero. The following proposition characterizes how incident frequency γ and mean incident duration $1/\theta$ affect the expected headway under the normal approximation.

PROPOSITION 10 (Impact of incidents on headway). Under Assumption 1 and the normal approximation, $\mathbb{E}[H_{Normal}^{(n)}]$ is nondecreasing in γ and $1/\theta$. It is strictly increasing in γ and, when $\gamma > 0$, strictly increasing in $1/\theta$.

Proposition 10 supports the comparative stability analysis in Section 4.4.

4.3.5. Distribution of passenger arrivals. We can now return to the passenger arrival distribution that enters the queueing model. Recall that $H_{\text{Normal}}^{(n)}$ is obtained by taking the positive part of the normal approximation to the raw headway $\widehat{H}^{(n)}$. Averaging the conditional Poisson PGF over this headway distribution gives the following result for $Y^{(n)}$.

PROPOSITION 11 (PGF of $Y^{(n)}$). *For each $n = 1, \dots, N$ and $\gamma > 0$, the PGF $Y(z)$ of $Y^{(n)}$ is*

$$Y(z) = \Phi\left(\frac{-\mu}{\sigma}\right) + e^{\mu\lambda^{(n)}(z-1) + \frac{\sigma^2(\lambda^{(n)}z - \lambda^{(n)})^2}{2}} \left[1 - \Phi\left(\frac{-\mu}{\sigma} - \sigma\lambda^{(n)}(z-1)\right)\right], \quad (38)$$

where $\mu = \overline{H} + \frac{2\gamma T^{(N)}}{\theta F}$ and $\sigma = \frac{2\sqrt{T^{(N)}}\gamma}{\theta}$ are the mean and standard deviation of the raw headway approximation $\widehat{H}_{\text{Normal}}^{(n)}$ respectively. When $\gamma = 0$, $Y^{(n)}$ is Poisson with mean $\lambda^{(n)}\overline{H}$ and $Y(z) = \exp[\lambda^{(n)}\overline{H}(z-1)]$.

Differentiating the PGF in Proposition 11 at $z = 1$ gives the first three factorial moments. We then convert these quantities to $\overline{Y}^{(n)}$, $\overline{\overline{Y}}^{(n)}$, and $\overline{\overline{\overline{Y}}}^{(n)}$. The first moment is shown below. The second and third central moments follow from the same differentiation and conversion, but their longer expressions are omitted for readability.

$$\overline{Y}^{(n)} = \left(\mu \cdot \Phi\left(\frac{\mu}{\sigma}\right) + \sigma \cdot \phi\left(\frac{\mu}{\sigma}\right)\right) \cdot \lambda^{(n)}. \quad (39)$$

When $\gamma = 0$, this expression is interpreted by continuity and gives $\overline{Y}^{(n)} = \lambda^{(n)}\overline{H}$.

4.3.6. Solving for the roots. With $Y(z)$ now specified, we turn to the C roots of $\mathcal{F}^{(n)}(z)$ in the closed unit disk that are required by the queue length and waiting time formulas. A nonlinear solver with a single initialization generally returns only one root, and repeated initializations may converge to the same location. The expression for $Y(z)$ makes this problem more difficult for the vehicle capacities considered here. We therefore develop a certified interpolation search that combines the structured initialization of Powell (1985) with deterministic interpolation and multiplicity certification. The algorithm supplies the required roots and reports a diagnostic failure when certification is unsuccessful.

Step 1. Initial root search. Following the structured angular initialization of Powell (1985), we place C initial values at equally spaced angles $2\pi(k + 1/2)/C$ for $k = 0, \dots, C - 1$. Let $\rho^{(n)} := \overline{Y}^{(n)} / \overline{S}^{(n)}$ denote utilization at station n . Their common radius is $1 - 0.5\rho^{(n)}$, bounded between 0.05 and 0.98. From each initial value, a local solver evaluates the real and imaginary parts of $\mathcal{F}^{(n)}(z)$

in Eq. 17. Candidates that pass the convergence, residual, and unit disk checks are retained together with their conjugates and grouped into distinct root clusters. Polar extrapolation from consecutive clusters provides additional initial values for the initial search.

The root locations often form an oval pattern. Figure 6 illustrates this pattern for different utilization ratios $\rho^{(n)}$ and available capacity distributions. As $\rho^{(n)}$ increases, the root locations generally move from a nearly circular pattern toward a more elongated pattern.

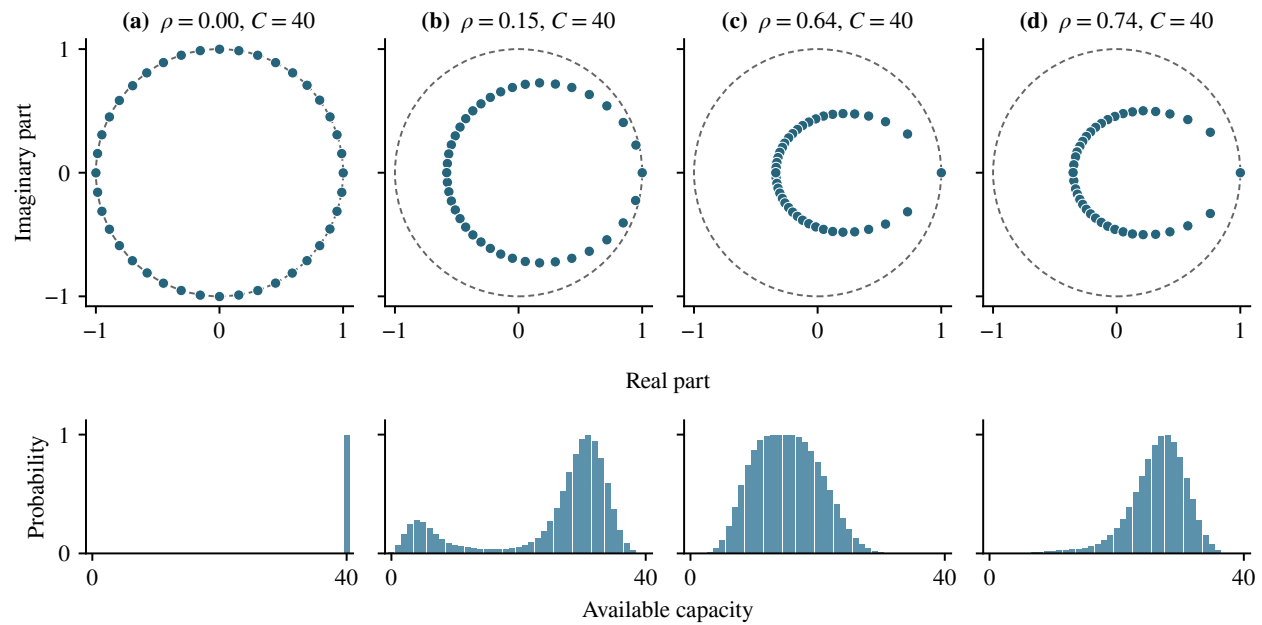


Figure 6 Characteristic roots and available capacity distributions across station utilization levels. Each column reports the certified roots and available capacity distribution for the displayed $\rho^{(n)}$ and C

Step 2. Interpolation search. The initial search may return fewer than C roots counted according to multiplicity. We order the distinct root clusters by complex argument and interpret adjacent pairs circularly. At interpolation level L , the algorithm interpolates both the modulus and the unwrapped argument of every adjacent pair at $L - 1$ locations. It places initial values at the interpolated modulus and at radial factors 0.85 and 1.15, with the modulus bounded by one. These radial values prevent a straight chord between two roots from moving all initial values toward the origin. The algorithm applies the local solver, refines each successful candidate, and updates the root clusters. If an iteration finds no new root location, it increases L . Because the coefficients of $\mathcal{F}^{(n)}(z)$ are real, the conjugate of every nonreal candidate is also evaluated. Figure 7 compares the distinct locations from the initial search with the root set after interpolation and certification.

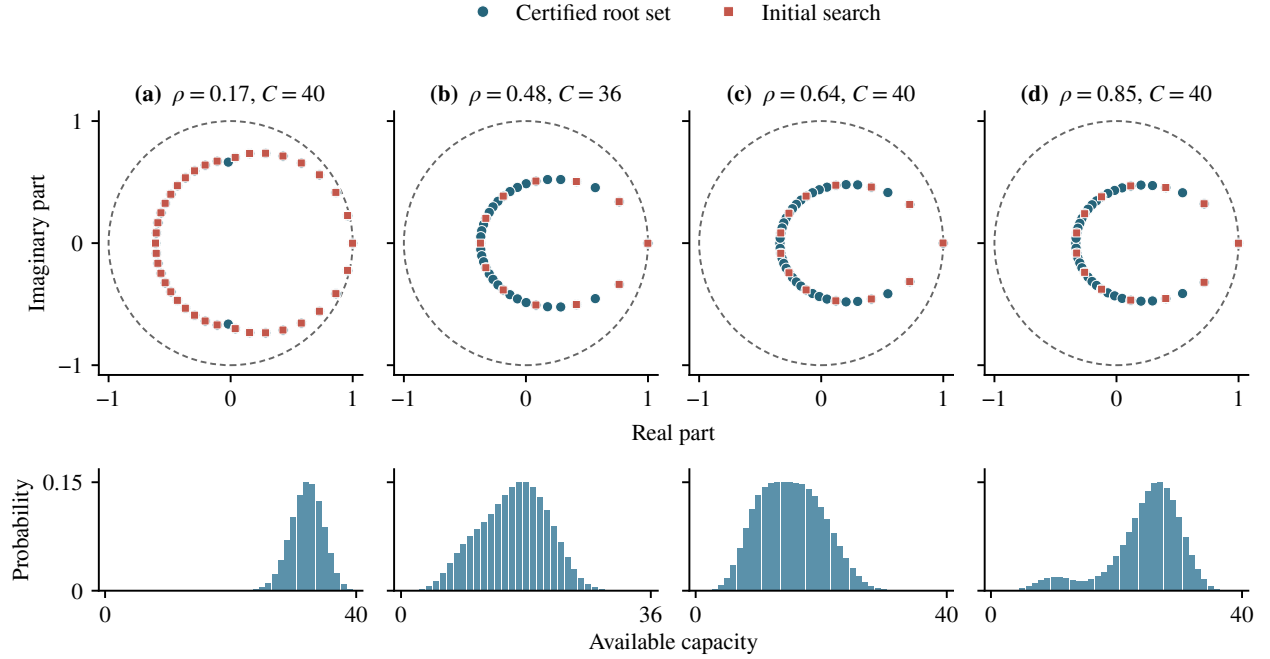


Figure 7 Comparison of the initial search and certified root set across station utilization and vehicle capacity settings

Step 3. Multiplicity certification and termination. Finding C distinct locations is not a valid stopping rule because a root may have multiplicity greater than one. Let $\mathcal{Z}^{(k)} := \{\zeta_0^{(k)}, \dots, \zeta_{M_k-1}^{(k)}\}$ contain the M_k distinct root clusters after iteration k . For cluster j , define the counterclockwise circular contour $\Gamma_j^{(k)} := \{\zeta_j^{(k)} + r_j^{(k)} \exp(\varphi\sqrt{-1}) \mid 0 \leq \varphi \leq 2\pi\}$. Its radius is smaller than half the distance to the nearest other cluster and no greater than 0.05. We use 0.05 when only one cluster is available. Thus, $\Gamma_j^{(k)}$ is tied to candidate $\zeta_j^{(k)}$ and separates it from other candidates. We use the contour only when $\mathcal{F}^{(n)}(z) \neq 0$ for every z on $\Gamma_j^{(k)}$.

$$m_j^{(k)} := \frac{1}{2\pi\sqrt{-1}} \int_{\Gamma_j^{(k)}} \frac{[\mathcal{F}^{(n)}]'(z)}{\mathcal{F}^{(n)}(z)} dz. \quad (40)$$

Because $\mathcal{F}^{(n)}(z)$ has no poles, $m_j^{(k)}$ equals the number of roots enclosed by $\Gamma_j^{(k)}$, counted according to multiplicity. It is also the winding number of $\mathcal{F}^{(n)}(z)$ around the origin along this contour. A value of one certifies a simple root in cluster j , while a value of two certifies a root of multiplicity two. We compute this value from the unwrapped change in the argument of $\mathcal{F}^{(n)}(z)$ and repeat the calculation with finer discretizations and smaller contours. A cluster is certified when these calculations return the same positive integer.

The algorithm terminates when every cluster is certified and the sum of the certified multiplicities equals C . Otherwise, it returns to Step 2 until a search limit is reached and then reports a diagnostic

failure. This certification follows numerical treatments of transform equations with possible multiple roots (Oblakova et al. 2019).

The complete procedure is provided in Algorithm EC.1 of the EC appendix.

4.4. Stability condition

The preceding transform calculations characterize the queue conditional on the existence of a steady state. We now establish when this steady state exists directly from the transient queue recursion. We call a station stable when its queue process admits a proper stationary probability distribution.

Assumption 3 separates the propagation of available capacity along the route from the evolution of the queue at an individual station. Upstream passenger flows determine $s^{(n)}$, so available capacity remains endogenous across stations. After $s^{(n)}$ is propagated to station n , the model uses its complete distribution to generate repeated service opportunities that are exogenous to the local queue, thereby carrying the effect of upstream crowding into the stationary single station analysis.

Assumption 3 is exact at the first station because every arriving vehicle has capacity C . Under persistent upstream saturation, the approximation is also reasonable because consecutive vehicles depart at capacity and independent binomial alighting determines the available capacity at station n . Outside these cases, independent renewal headways generate new passenger arrivals for successive vehicles, while alighting at intermediate stations weakens the persistence of upstream loads. These mechanisms limit dependence in available capacity when suspensions are short and rarely affect several vehicle runs. When long disruptions or common operating conditions sustain correlated vehicle loads, the approximation becomes weaker. Section 5 assesses it through recursive simulation.

PROPOSITION 12 (Stability condition). *Under Assumptions 1 and 3 and the normal approximation in Proposition 9, the bulk service queueing system at station n with $\bar{S}^{(n)} > 0$ is stable if and only if*

$$\rho^{(n)} = \frac{\bar{Y}^{(n)}}{\bar{S}^{(n)}} = \frac{(\mu \cdot \Phi(\frac{\mu}{\sigma}) + \sigma \cdot \phi(\frac{-\mu}{\sigma})) \cdot \lambda^{(n)}}{\sum_{u=0}^C s_u^{(n)} u} = \frac{\lambda^{(n)} \cdot \mathbb{E}[H_{Normal}^{(n)}]}{\sum_{u=0}^C s_u^{(n)} u} < 1, \quad (41)$$

where $\rho^{(n)}$ is the utilization ratio at station n . When $\gamma = 0$, the expression is interpreted by continuity, giving $\mathbb{E}[H_{Normal}^{(n)}] = \bar{H}$ and $\rho^{(n)} = \lambda^{(n)} \bar{H} / \sum_{u=0}^C s_u^{(n)} u$.

The proof in the EC appendix writes the residual queue as a Lindley recursion (Lindley 1952) and applies the stability result of Loynes (1962). It therefore establishes the existence of the stationary distribution without using $q_0^{(n)}$ or any other quantity derived under stationarity.

Proposition 12 states that station n is stable when the mean number of passenger arrivals within a headway is smaller than the mean available capacity of an arriving vehicle after alighting. Proposition 10 shows that a higher incident rate or a longer mean incident duration increases $\mathbb{E}[H_{\text{Normal}}^{(n)}]$ under the conditions stated there, making instability more likely. The result therefore quantifies the throughput loss induced by incidents.

REMARK 3. The distribution $s^{(n)}$ depends on the upstream queue and vehicle load distributions, which in turn require the characteristic roots at stations $1, \dots, n-1$. Stability at station n must therefore be evaluated after propagating the preceding stations. At station 1, $s_C^{(1)} = 1$ and $s_u^{(1)} = 0$ for $u = 0, \dots, C-1$, so Eq. 41 gives $\rho^{(1)} = \lambda^{(1)} \mathbb{E}[H_{\text{Normal}}^{(1)}]/C$ for a direct stability assessment.

REMARK 4. Proposition 12 concerns station level stability. A route is stable if and only if $\rho^{(n)} < 1$ for every $n = 1, \dots, N$.

REMARK 5. If station $n-1$ is stable, its departing load distribution determines $s^{(n)}$ through Section 4.1, after which stability at station n can be evaluated. If station $n-1$ is unstable, station n may nevertheless be stable because passengers can alight there. In this case, vehicles depart station $n-1$ at capacity, so $v_C^{(n-1)} = 1$ and $v_k^{(n-1)} = 0$ for $k = 0, \dots, C-1$. Binomial alighting then gives $\bar{S}^{(n)} = \alpha^{(n)} C$, and the stability condition becomes $\rho^{(n)} = \lambda^{(n)} \mathbb{E}[H_{\text{Normal}}^{(n)}]/[\alpha^{(n)} C] < 1$.

4.5. Summary of calculation procedure

Algorithm 1 combines the preceding analytical results into a sequential calculation over the N stations of the route.

Algorithm 1 Calculation of performance indicators

```

1: Initialize  $v_0^{(0)} \leftarrow 1$  and  $v_k^{(0)} \leftarrow 0$  for  $k = 1, \dots, C$ 
2: for  $n = 1 : N$  do
3:    $g^{(n)} = v^{(n-1)} A^{(n)}$  ▷ Proposition 1
4:    $s_k^{(n)} \leftarrow g_{C-k}^{(n)}$  for  $k = 0, 1, \dots, C$  ▷ Eq. 12
5:   Calculate  $\bar{S}^{(n)}$ ,  $\bar{\bar{S}}^{(n)}$ , and  $\bar{\bar{\bar{S}}}^{(n)}$  based on  $s^{(n)}$ .
6:   Calculate  $\bar{Y}^{(n)}$ ,  $\bar{\bar{Y}}^{(n)}$ , and  $\bar{\bar{\bar{Y}}}^{(n)}$  ▷ Section 4.3.5
7:   if  $\bar{Y}^{(n)} < \bar{S}^{(n)}$  then ▷ Station  $n$  is stable
8:     Solve for the roots  $z_0^*, \dots, z_{C-1}^*$  of the denominator of  $Q(z)$  in Eq. 16 ▷ Section 4.3.6
9:     Calculate  $q_0^{(n)}, \dots, q_{C-1}^{(n)}$  from  $z_0^*, \dots, z_{C-1}^*$  ▷ Section 4.3.2
10:    Calculate  $\mathbb{E}[Q^{(n)}]$  and  $\text{Var}[Q^{(n)}]$ , and, if  $\lambda^{(n)} > 0$ , calculate  $\mathbb{E}[W^{(n)}]$  and  $\text{Var}[W^{(n)}]$  ▷ Eqs. 22–25
11:     $v^{(n)} = g^{(n)} B^{(n)}$  ▷ Eq. 14.  $B^{(n)}$  is a function of  $q_{0:C-1}^{(n)}$ 
12:  else ▷ Station  $n$  is not stable
13:    Set  $\mathbb{E}[Q^{(n)}]$ ,  $\text{Var}[Q^{(n)}]$ ,  $\mathbb{E}[W^{(n)}]$ , and  $\text{Var}[W^{(n)}]$  to infinity
14:    Set  $v_C^{(n)} \leftarrow 1$  and  $v_k^{(n)} \leftarrow 0$  for  $k = 0, 1, \dots, C-1$ 

```

5. Numerical experiments

This section illustrates the analytical framework on a bus route with ten stations and assesses the resulting resilience indicators against recursive FIFO simulation. The experiments isolate the effects of vehicle capacity, incident frequency, incident duration, scheduled headway, and demand, with particular attention to stations with limited capacity slack. All analytical results use the certified interpolation root search described in Section 4.3.6. Section EC.2 of the EC appendix reports the numerical settings used for the search.

5.1. Experimental design

We use an example bus route adapted from Islam et al. (2015). Table 1 reports the attributes of its ten stations. Figure 8 shows the route layout. Travel time without incidents is 5 minutes between consecutive stations, total cycle time is $\bar{E} = 100$ minutes, and travel time from the hub to the last station is $T^{(N)} = 50$ minutes.

Table 1 Example bus system parameters. $\lambda^{(n)}$ and $\alpha^{(n)}$ are the passenger arrival rate and alighting probability at station n

Station	$\lambda^{(n)}$ (passengers/min)	$\alpha^{(n)}$	Station	$\lambda^{(n)}$ (passengers/min)	$\alpha^{(n)}$
1	0.75	0	6	1	0.8
2	2	0	7	0.75	0.5
3	0.75	0.1	8	1.5	0.3
4	0.5	0.25	9	0.2	0.75
5	0.5	0.25	10	0	1

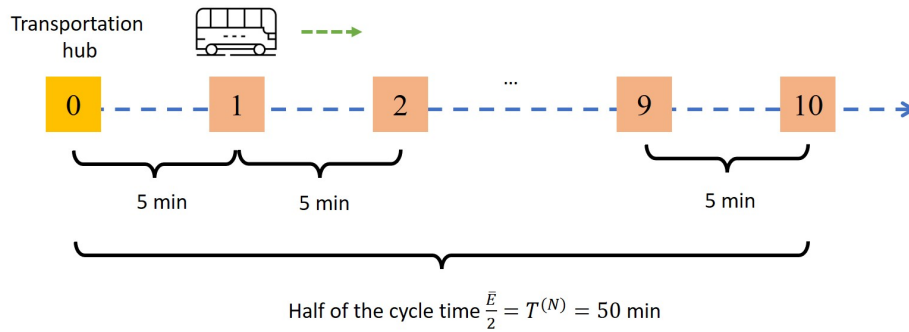


Figure 8 Case study route layout

To test the sensitivity of resilience indicators to different parameters, we consider different values of C , θ , γ , $\bar{H}$, and demand in Table 2. The demand factor multiplies the base arrival rates $\lambda^{(n)}$ in Table 1. Keeping the route cycle time at $\bar{E} = 100$ minutes, we select the integer fleet sizes

$\bar{F} \in \{50, 25, 14\}$, which yield $\bar{H} = \bar{E}/\bar{F} \in \{2, 4, 50/7\}$ minutes. The reference scenario sets $\bar{F} = 25$ and $\bar{H} = 4$ minutes; each test varies one parameter and retains all other reference values.

Table 2 Scenario design for vehicle capacity C , incident rate γ , incident duration rate θ , scheduled headway $\bar{H}$, and demand factor

Parameters	Tested values	Reference value
C	$\{30, 34, 38\}$	34
γ (/min)	$\{0, 1/10, 1/5, 1/3\}$	1/5
θ (/min)	$\{2, 1, 1/2\}$	1
$\bar{H}$ (min)	$\{2, 4, 50/7\}$	4
Demand factor	$\{0.5, 0.75, 1\}$	0.75

5.2. Performance indicators

Figure 9 shows the mean and standard deviation of queue length under the tested scenarios. The patterns are consistent with the passenger arrival and alighting rates. Expected queues are relatively high at stations 2 and 8, whereas the expected queue at station 10 is zero because $\lambda^{(10)} = 0$.

Figure 9a shows limited sensitivity to vehicle capacity within the tested range because capacity is not fully utilized under the reference setting. Figure 9b shows the effect of incident occurrence rate γ . When $\gamma = 0$, the expected queue length at station 8 is 4.5 passengers. As γ increases, expected queue lengths and their standard deviations rise. At $\gamma = 1/3$ per minute, the expected queue at station 8 reaches 7.2 passengers. Incident duration produces a similar pattern in Figure 9c. When mean incident duration is 30 seconds and $\theta = 2$, $\mathbb{E}[Q^{(8)}] = 5.0$. When mean incident duration is 2 minutes and $\theta = 1/2$, this expectation increases to 9.3. Both effects are larger at crowded stations.

Figure 9d shows that a longer scheduled headway lowers the service rate and increases the expected queue. As $\bar{H}$ increases from 2 minutes to 50/7 minutes, the expected queue at station 8 increases from 4.1 to 9.9 passengers. Figure 9e shows a similar pattern for demand. Increasing the demand factor from 0.5 to 1.0 raises the expected queue at station 8 from 4.1 to 8.2 passengers. The effects of $\bar{H}$ and demand are more uniform across stations than the effects of incident frequency and duration.

Figure 10 shows the mean and standard deviation of passenger waiting time. Station 10 is omitted because its zero passenger arrival rate produces no waiting time observations. Downstream stations generally have higher waiting time means and standard deviations because headway variability accumulates along the route. At congested stations such as stations 3 and 8, capacity constraints leave some passengers behind and further increase waiting time.

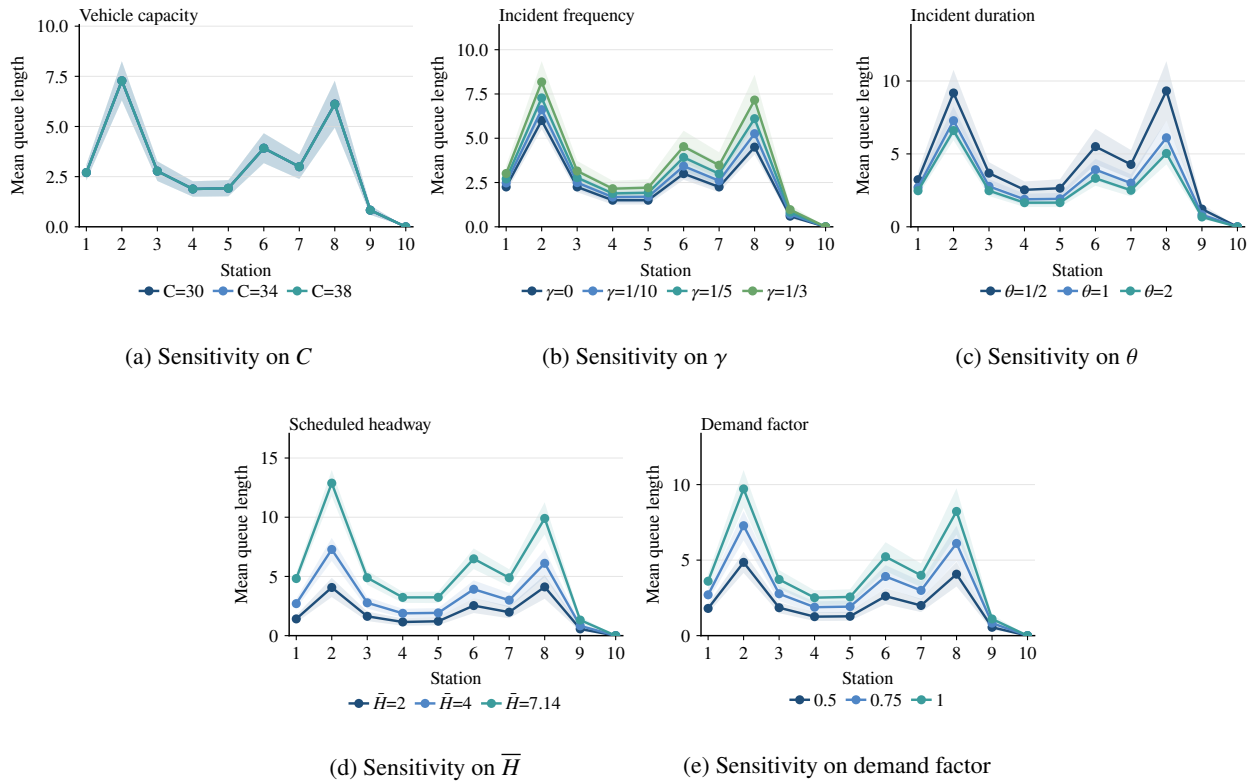


Figure 9 Mean queue length across stations. Lines show the means, and shaded bands extend 0.2 standard deviations above and below each mean

Figure 10a shows that capacity has a limited effect on waiting time within the tested range. Figures 10b and 10c show the effects of γ and θ . Higher values of γ and $1/\theta$ increase the expected headway and, in turn, mean waiting time at stations with passenger demand, particularly at crowded stations. When $\gamma = 0$, there are no incidents, left behind passengers, or headway irregularities. Expected waiting time at stations 1 through 9 is then $\bar{H}/2 = 2$ minutes under the reference headway of 4 minutes. When γ increases to $1/5$, station 3 has left behind passengers and its expected waiting time increases to 3.5 minutes. When θ decreases from 2 to $1/2$, expected waiting time at station 8 increases from 3.0 to 8.8 minutes.

A larger scheduled headway directly increases waiting time at stations with passenger demand, and passengers are left behind at stations 3 and 8 when $\bar{H} = 50/7$ minutes. Demand does not alter the headway distribution, so its effect on waiting time remains limited until capacity constraints leave passengers behind. At station 8, increasing the demand factor from 0.5 to 1.0 raises expected waiting time only from 4.77 to 4.83 minutes within the tested range.

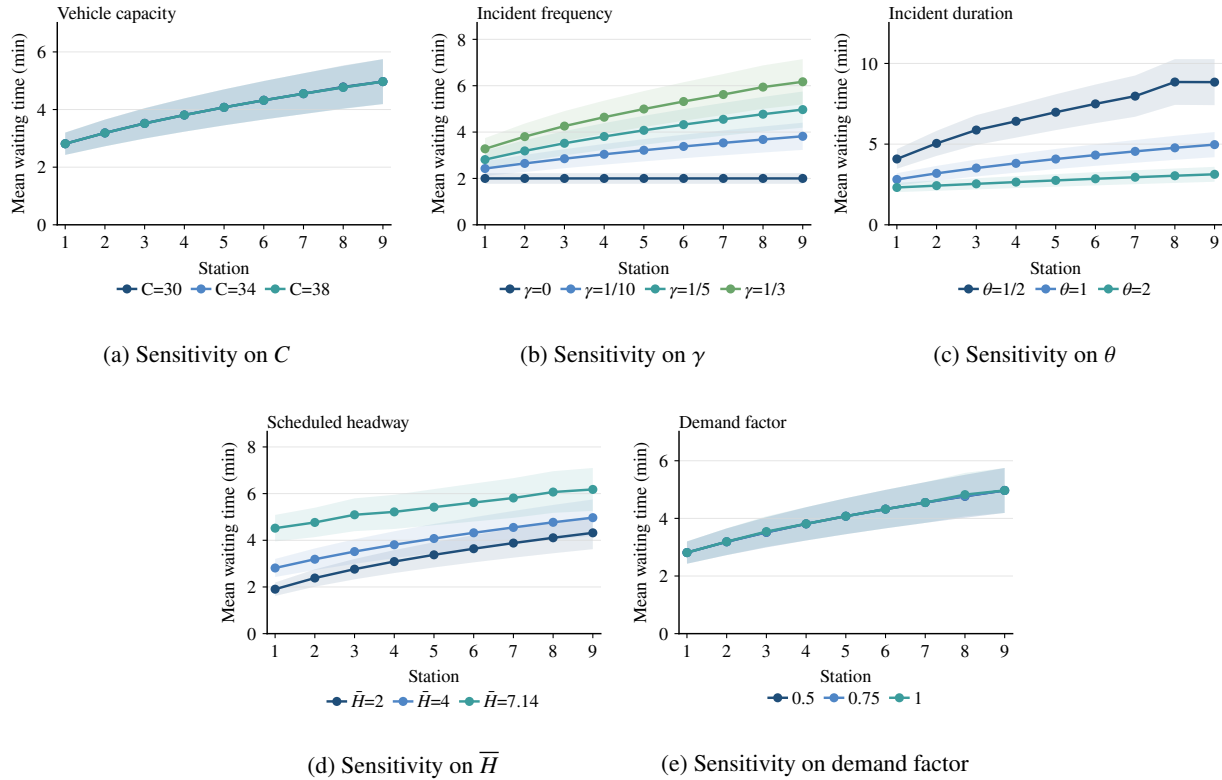


Figure 10 Mean waiting time across stations. Lines show the means, and shaded bands extend 0.2 standard deviations above and below each mean

5.3. Comparison between simulation and analytical results

To assess the renewal headway, marginal route decoupling, and normal distribution approximations, we develop a recursive FIFO simulation model for the expectation and variance of queue length and waiting time. Unlike the analytical renewal model, the simulation carries the effect of a delayed vehicle to its successors through realized departure times and jointly evolves passenger queues and vehicle loads; see Section EC.4 of the EC appendix.

The assessment covers seven settings from Table 2, including the reference setting and variations in capacity, headway, demand, incident frequency, and incident duration. For each setting, we conduct 20 independent replications of 10,000 vehicle runs using distinct random seeds and exclude the first 10% of vehicle runs. Figure 11 compares the analytical results with the means across replications for the reference setting. The whiskers give 95% confidence intervals (CIs) based on variation across replications. Station 10 has $\lambda^{(10)} = 0$, so no passenger arrives there and no waiting time observation exists. We therefore report waiting time through station 9. All error calculations also use stations 1 through 9 so that the four performance measures have common support.

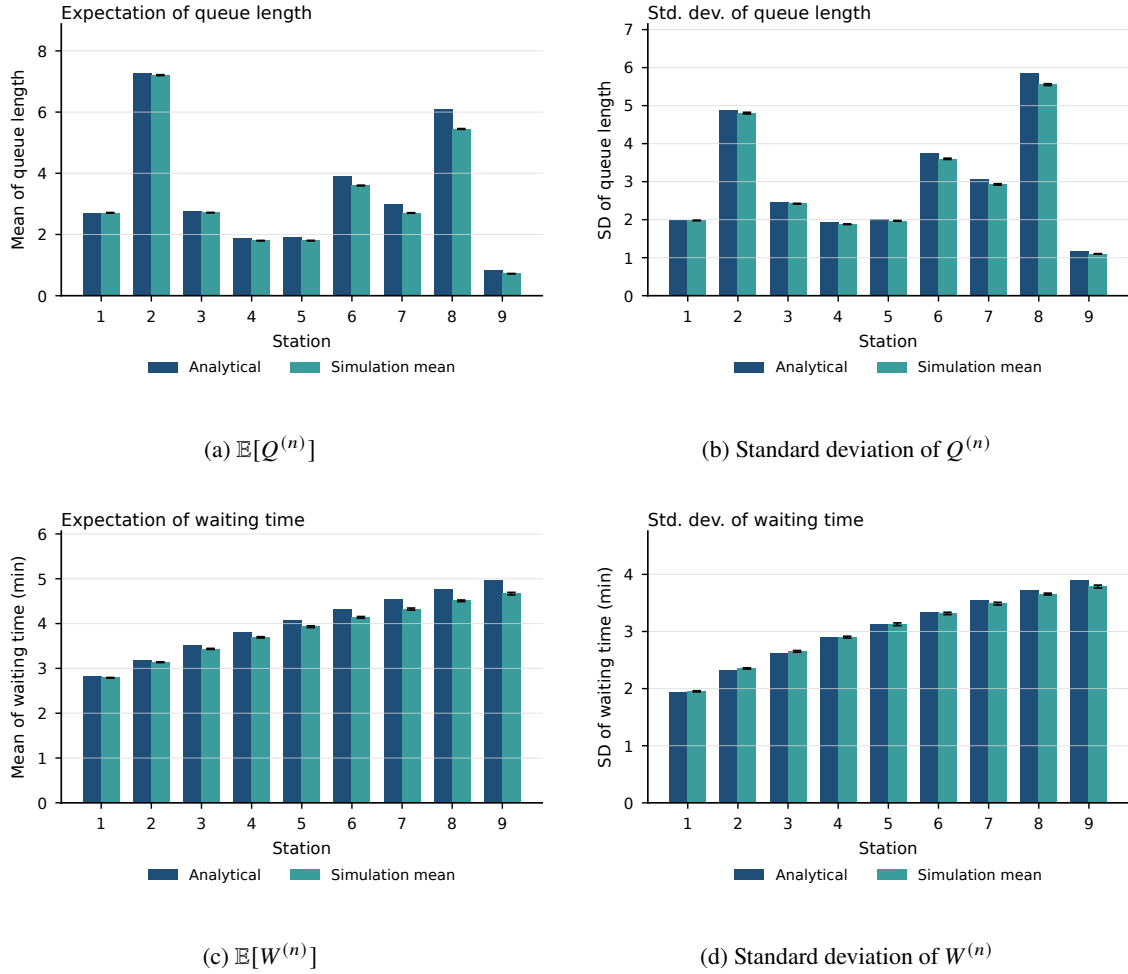


Figure 11 Comparison between simulation and analytical results for the reference setting.

Table 3 reports the relative ℓ_1 error in percentage terms and the average half width (HW) of the station CIs. For a given performance metric, $x_{\text{Sim}}^{(n)}$ denotes the mean estimate across the 20 independent simulation replications at station n , and $x_{\text{Ana}}^{(n)}$ denotes the corresponding result obtained from the analytical model. The relative ℓ_1 error is calculated as $100 \sum_{n=1}^9 |x_{\text{Sim}}^{(n)} - x_{\text{Ana}}^{(n)}| / \sum_{n=1}^9 |x_{\text{Ana}}^{(n)}|$. Under the reference setting, the queue length errors are 5.64% for the mean and 3.27% for the standard deviation. The corresponding waiting time errors are 3.87% and 1.26%.

Across all tested settings, the waiting time errors remain below 8.21%. The largest discrepancy is the 16.58% error in mean queue length when the mean incident duration is 2 minutes. This setting is also where the short suspension premise and the independent renewal approximation are most strained. The average confidence interval half width over the seven settings is at most 0.013 passengers for mean queue length, 0.025 passengers for its standard deviation, 0.039 minutes

for mean waiting time, and 0.052 minutes for its standard deviation. The sampling uncertainty is therefore small relative to the main analytical differences.

Table 3 Relative ℓ_1 errors and confidence interval half widths for queue length and waiting time

Setting	$\mathbb{E}[Q^{(n)}]$		$\text{SD}[Q^{(n)}]$		$\mathbb{E}[W^{(n)}]$		$\text{SD}[W^{(n)}]$	
	Error (%)	CI HW	Error (%)	CI HW	Error (%)	CI HW	Error (%)	CI HW
$C = 30$	5.40	0.008	3.34	0.012	3.88	0.016	1.32	0.020
Reference	5.64	0.009	3.27	0.013	3.87	0.016	1.26	0.018
$\bar{H} = 50/7$	0.67	0.012	1.49	0.015	0.93	0.013	0.31	0.016
Demand factor = 1	4.89	0.010	3.53	0.016	4.04	0.014	1.38	0.016
$\gamma = 0$	0.11	0.008	0.14	0.006	0.05	0.004	0.07	0.002
$\gamma = 1/3$	7.97	0.010	4.85	0.016	5.48	0.020	2.68	0.024
$\theta = 1/2$	16.58	0.011	8.68	0.025	8.20	0.039	3.65	0.052

The confidence interval half widths are averaged across stations 1 through 9. Queue length is measured in passengers, and waiting time is measured in minutes. In each nonreference row, only the listed parameter differs from its reference value.

The results provide numerical evidence for the accuracy of the approximations within the tested parameter range, but they do not establish exact equivalence between the analytical renewal process and the recursive FIFO process. One source of discrepancy is the normal approximation of the raw headway distribution. As shown in Figure 5, the compound Poisson based raw headway concentrates more probability density near the mean than the normal approximation and therefore assigns a lower probability to large deviations from the scheduled headway. A second source is the independent redraw of each analytical headway, whereas the recursive FIFO simulation carries delay forward.

6. Conclusion and discussion

This paper proposes a stochastic framework for quantifying the resilience of PTSs under short random service suspensions. Specifically, we introduce an independent renewal approximation for nonnegative headways and derive stability conditions and closed form expressions for the mean and variance of queue length and waiting time at each station. The common headway marginal preserves the incident effects implied by a two state vehicle speed process, while the renewal and route decoupling assumptions make the bulk service analysis tractable. The stability conditions imply that the system is more likely to be unstable under high incident rates, long incident durations, high demand, low service frequency, and low vehicle capacity. We implement the model on an example bus route adapted from the literature, and the sensitivity analysis shows that congested stations are more vulnerable to random service suspensions. Comparison with a recursive FIFO simulation across seven parameter settings quantifies the approximation error and shows that the discrepancy increases when incident durations move beyond the intended short suspension setting.

The proposed model can support public transit planning, operations, and management in several ways. For the design and planning of resilient systems under random system interruptions, the model can inform headway design and vehicle capacity determination. Its queue length estimates can be used to evaluate the layout and capacity of congested stations. Historical demand and incident information can be used with the model to monitor system resilience and identify critical stations. The estimates of waiting time and queue length can also support cost benefit analyses of service improvements. For example, the model can identify whether a larger vehicle, shorter headway, or higher maintenance frequency is the most cost effective way to keep waiting time below a threshold.

Several extensions are possible. A network level model would differ primarily through the inclusion of transfer passengers. Transfer demand could enter the arrival process, but the model would require additional parameters related to transfers and alighting rates. The renewal approximation could also be replaced by a dependence aware model that carries recursive FIFO delay and the joint passenger queue and vehicle load state. Another extension would consider the partial interruptions studied in Section 2.2. Vehicles retain a positive speed during a partial interruption, which requires a revised headway distribution. A final extension would relax the assumption of no passenger balking or renegeing and incorporate more complicated passenger behavior.

Acknowledgments

This work was in part supported by SJTU UM Joint Institute, J. Wu & J. Sun Endowment Fund, C2SMART University Transportation Center, and NYU Tandon School of Engineering.

References

- Adan IJBF, van Leeuwen JSH, Winands EMM, 2006 *On the application of Rouché's theorem in queueing theory. Operations Research Letters* 34(3):355–360.
- Bailey NT, 1954 *On queueing processes with bulk service. Journal of the Royal Statistical Society: Series B (Methodological)* 16(1):80–87.
- Bellei G, Gkoumas K, 2010 *Transit vehicles' headway distribution and service irregularity. Public transport* 2(4):269–289.
- Chaudhry ML, Madill B, Briere G, 1987 *Computational analysis of steady-state probabilities of $m/g a, b/1$ and related nonbulk queues. Queueing systems* 2(2):93–114.
- Daganzo CF, 2009 *A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons. Transportation Research Part B: Methodological* 43(10):913–921.
- Hickman MD, 2001 *An analytic stochastic model for the transit vehicle holding problem. Transportation Science* 35(3):215–237.

- Islam MK, Vandebona U, Dixit VV, Sharma A, 2014 *A bulk queue model for the evaluation of impact of headway variations and passenger waiting behavior on public transit performance. IEEE Transactions on Intelligent Transportation Systems* 15(6):2432–2442.
- Islam MK, Vandebona U, Dixit VV, Sharma A, 2015 *A model to evaluate the impact of headway variation and vehicle size on the reliability of public transit. IEEE Transactions on Intelligent Transportation Systems* 16(4):1840–1850.
- Jin JG, Teo KM, Odoni AR, 2016 *Optimizing bus bridging services in response to disruptions of urban transit rail networks. Transportation Science* 50(3):790–804.
- Krishnamoorthy A, Pramod PK, Chakravarthy SR, 2014 *Queues with interruptions: a survey. Top* 22(1):290–320.
- Lindley DV, 1952 *The theory of queues with a single server. Mathematical Proceedings of the Cambridge Philosophical Society* 48(2):277–289, URL <http://dx.doi.org/10.1017/S0305004100027638>.
- Loynes RM, 1962 *The stability of a queue with non-independent inter-arrival and service times. Mathematical Proceedings of the Cambridge Philosophical Society* 58(3):497–520, URL <http://dx.doi.org/10.1017/S0305004100036781>.
- Madan K, 1989 *A single channel queue with bulk service subject to interruptions. Microelectronics Reliability* 29(5):813–818.
- Mo B, Koutsopoulos HN, Shen ZJM, Zhao J, 2023 *Robust path recommendations during public transit disruptions under demand uncertainty. Transportation Research Part B: Methodological* 169:82–107.
- Mo B, Koutsopoulos HN, Shen ZJM, Zhao J, 2025 *Individual path recommendation under public transit service disruptions considering behavior uncertainty. Transportation Science* 59(6):1235–1258.
- Mo B, Ma Z, Koutsopoulos HN, Zhao J, 2020 *Capacity-constrained network performance model for urban rail systems. Transportation Research Record* 2674(5):59–69.
- Mo B, Von Franque MY, Koutsopoulos HN, Attanucci JP, Zhao J, 2022 *Impact of unplanned long-term service disruptions on urban public transit systems. IEEE Open Journal of Intelligent Transportation Systems* 3:551–569.
- Oblakova A, Al Hanbali A, Boucherie RJ, van Ommeren JKW, Zijm WHM, 2019 *An exact root-free method for the expected queue length for a class of discrete-time queueing systems. Queueing Systems* 92(3–4):257–292, URL <http://dx.doi.org/10.1007/s11134-019-09614-1>.
- Osuna E, Newell GF, 1972 *Control strategies for an idealized public transportation system. Transportation Science* 6(1):52–72.
- Powell WB, 1981 *Stochastic delays in transportation terminals: New results in the theory and application of bulk queues*. Ph.D. thesis, Massachusetts Institute of Technology.
- Powell WB, 1985 *Analysis of vehicle holding and cancellation strategies in bulk arrival, bulk service queues. Transportation Science* 19(4):352–377.
- Rahimi E, Shamshiripour A, Shabanpour R, Mohammadian A, Auld J, 2019 *Analysis of transit users' waiting tolerance in response to unplanned service disruptions. Transportation Research Part D: Transport and Environment* 77:639–653.

Electronic Companion

EC.1. Modeling assumptions

Table EC.1 summarizes route scope and stochastic primitives, and Table EC.2 summarizes dependence and service sequence approximations.

Table EC.1 Route scope and stochastic primitives

Assumption	Operational interpretation
Single route with ordered stations	The model targets line level resilience for one bus route or one rail direction, where vehicles serve stops in a fixed order. This is the natural unit for evaluating station queues, vehicle capacity propagation, and headway disruptions before adding transfer interactions.
Poisson passenger arrivals	Passenger arrivals are treated as random at the aggregate stop level during the analysis period. This is appropriate when individual arrivals are not perfectly synchronized with vehicle departures and when the arrival rate $\lambda^{(n)}$ can be estimated from observed demand.
Binomial alighting	Each onboard passenger has a station specific probability of alighting. This provides a parsimonious way to propagate vehicle load and downstream available capacity when only aggregate alighting shares or proportions by origin and destination are available.
Negligible dwell time	The analysis focuses on headway changes caused by short service suspensions rather than passenger service time at stops. When dwell time is small relative to the operating headway, passenger arrivals during dwell are second order, and systematic dwell components can be absorbed into the effective headway if needed.
No balking or reneging	The paper studies short suspensions, during which waiting passengers are more likely to remain in the system than abandon their trip. The assumption therefore keeps the resilience indicators focused on capacity, headway, and passengers left behind, without modeling route choice behavior during longer disruptions.
Full stop failure state	A full stop is a conservative first order representation of incidents such as signal failures, vehicle breakdowns, or blocked movements. Partial slowdowns can be approximated by an effective suspension duration or by extending the two state speed process.

Table EC.2 Dependence and service sequence approximations

Assumption	Operational interpretation
No overtaking	Vehicles serve each station in dispatch order. This rule directly represents rail lines where passing is infeasible and bus services where passing is prohibited or uncommon. The recursive first in, first out (FIFO) model holds a vehicle when its free departure time precedes the realized departure time of the preceding vehicle.
Independent exogenous renewal headways	At each station, headways are independent across vehicle runs and independent of prior passenger queue and vehicle load states. Their common marginal law retains the disruption effect over one headway implied by two independent incident duration draws. Serial dependence and delay propagation across multiple vehicles are omitted. The approximation is most appropriate when holding or dispatch control resets the next service interval from the current departure rather than recovering the absolute timetable. It is assessed against a recursive FIFO simulation.
Marginal route decoupling	The assumption is exact at the first station. At downstream stations, separate local demand, independent renewal headways, and binomial alighting weaken dependence between the remaining onboard load and local queue when suspensions are short and stations are stable. The simulation assesses the approximation.
Exogenous station service sequence	Once the marginal available capacity distribution has been propagated from upstream, successive service opportunities at a station are independent draws from that distribution and are independent of the local queue and arrival processes. The complete capacity distribution carries the effect of upstream crowding into the station model, whereas the analytical queue omits temporal dependence in capacity.
Nonnegative headways without later recovery	Departure epochs are accumulated from nonnegative headways, so the analytical service sequence is nondecreasing by construction. A nonpositive raw headway becomes a zero headway, representing bunching, and the lost separation is not recovered in a later headway. The resulting analytical service sequence follows the renewal representation and differs from the realized FIFO trajectory.

EC.2. Certified interpolation root search

Algorithm [EC.1](#) gives the complete procedure for the three search steps described in Section [4.3.6](#) and reports the parameter settings used for every station and numerical scenario.

Algorithm EC.1 Certified interpolation search

Require: Characteristic function $\mathcal{F}^{(n)}$, passenger arrival probability generating function $Y(z)$, capacity distribution $s^{(n)}$, capacity C , and utilization $\rho^{(n)} = \bar{Y}^{(n)} / \bar{S}^{(n)}$

- 1: Set the residual tolerance $\tau_{\text{res}} \leftarrow 10^{-9}$ and clustering tolerance $\tau_{\text{dup}} \leftarrow 10^{-4}$
- 2: Set the unit disk tolerance $\tau_{\text{disk}} \leftarrow 10^{-8}$
- 3: Set $K_{\text{max}} \leftarrow \max\{12, C\}$ and $L_{\text{max}} \leftarrow \max\{4, \lceil C/2 \rceil\}$
- 4: Define the scaled residual

$$\text{Res}^{(n)}(z) := \frac{|\mathcal{F}^{(n)}(z)|}{|z|^C + |Y(z) \sum_{u=0}^C s_u^{(n)} z^{C-u}|}$$
- 5: For a candidate set C , define $E_{\text{dup}}(C) := \{\{z, z'\} \subseteq C : z \neq z', |z - z'| \leq \tau_{\text{dup}}\}$
The connected components of $(C, E_{\text{dup}}(C))$ are the root clusters
- 6: Set $r_0 \leftarrow \min\{0.98, \max\{0.05, 1 - 0.5\rho^{(n)}\}\}$
- 7: Initialize the candidate set $C^{(0)} \leftarrow \{1\}$
- 8: **for** $k = 0, \dots, C - 1$ **do**
- 9: Set $z^{\text{Ini}} \leftarrow r_0 \exp[2\pi(k + 1/2)\sqrt{-1}/C]$ following [Powell \(1985\)](#)
- 10: Apply the local solver to $\mathcal{F}^{(n)}(z) = 0$ from z^{Ini}
- 11: **if** the solver returns z with $\text{Res}^{(n)}(z) \leq \tau_{\text{res}}$ and $|z| \leq 1 + \tau_{\text{disk}}$ **then**
- 12: Add z and its conjugate to $C^{(0)}$
- 13: Form the root clusters from $C^{(0)}$ using the graph rule above and apply polar extrapolation
- 14: Certify the cluster multiplicities using Eq. [40](#); obtain $\mathcal{Z}^{(0)}$, M_0 , and the total certified multiplicity N_0
- 15: Set $L \leftarrow 2$ and $k \leftarrow 0$
- 16: **while** ($N_k < C$ or at least one cluster is uncertified) and $k < K_{\text{max}}$ and $L \leq L_{\text{max}}$ **do**
- 17: Order $\mathcal{Z}^{(k)} = \{\zeta_0^{(k)}, \dots, \zeta_{M_k-1}^{(k)}\}$ by complex argument
- 18: Initialize $\mathcal{Z}^{\text{Ini}} \leftarrow \emptyset$
- 19: **for** $i = 0, \dots, M_k - 1$ **do**
- 20: Set $j \leftarrow (i + 1) \bmod M_k$
- 21: Set $r_i \leftarrow |\zeta_i^{(k)}|$ and $r_j \leftarrow |\zeta_j^{(k)}|$; unwrap $\phi_i \leftarrow \arg(\zeta_i^{(k)})$ and $\phi_j \leftarrow \arg(\zeta_j^{(k)})$ so that $\phi_j > \phi_i$
- 22: **for** $d = 1, \dots, L - 1$ **do**
- 23: Set $r_{i,d} \leftarrow (1 - d/L)r_i + (d/L)r_j$ and $\phi_{i,d} \leftarrow (1 - d/L)\phi_i + (d/L)\phi_j$
- 24: **for** $a \in \{0.85, 1, 1.15\}$ **do**
- 25: Add $\min\{1, ar_{i,d}\} \exp(\phi_{i,d}\sqrt{-1})$ to $\mathcal{Z}^{\text{Ini}}$
- 26: Apply the local solver from every initial value in $\mathcal{Z}^{\text{Ini}}$ and refine every successful output
- 27: Collect each output and its conjugate with $\mathcal{Z}^{(k)}$ in $C^{(k+1)}$
- 28: $\tilde{C}^{(k+1)} \leftarrow \{z \in C^{(k+1)} : \text{Res}^{(n)}(z) \leq \tau_{\text{res}}, |z| \leq 1 + \tau_{\text{disk}}\}$
- 29: Let $C_0^{(k+1)}, \dots, C_{M_{k+1}-1}^{(k+1)}$ be the root clusters obtained from $\tilde{C}^{(k+1)}$ using the graph rule above
- 30: Set $\zeta_j^{(k+1)} \leftarrow |C_j^{(k+1)}|^{-1} \sum_{z \in C_j^{(k+1)}} z$ for each cluster; certify its multiplicity and obtain $\mathcal{Z}^{(k+1)}$, M_{k+1} , and N_{k+1}
- 31: Remove an uncertified cluster only if the scaled residual of its center exceeds $\max\{10^{-13}, \tau_{\text{res}}/100\}$
- 32: **if** $M_{k+1} = M_k$ **then**
- 33: Set $L \leftarrow L + 1$
- 34: Set $k \leftarrow k + 1$
- 35: **if** $N_k = C$ and every cluster is certified **then**
- 36: Return every cluster center repeated according to its certified multiplicity
- 37: **else**
- 38: Return a diagnostic failure with N_k , the largest residual, k , and L

EC.3. Proofs

EC.3.1. Proof of Proposition 1

Suppose that vehicle l carries i passengers when it reaches station n . During alighting, each passenger independently remains onboard with probability $1 - \alpha^{(n)}$. The remaining load therefore follows a binomial distribution with parameters i and $1 - \alpha^{(n)}$. For every $j = 0, \dots, i$, the corresponding conditional probability is $a_{ij}^{(n)} = \binom{i}{j} (1 - \alpha^{(n)})^j (\alpha^{(n)})^{i-j}$, as stated in Eq. 13.

The law of total probability gives $g_j^{(n,l)} = \sum_{i=0}^C v_i^{(n-1,l)} a_{ij}^{(n)}$. In vector form, this relation is

$$g^{(n,l)} = v^{(n-1,l)} A^{(n)}, \quad (\text{EC.1})$$

where $v^{(n-1,l)} = [v_0^{(n-1,l)}, \dots, v_C^{(n-1,l)}]$ and $g^{(n,l)} = [g_0^{(n,l)}, \dots, g_C^{(n,l)}]$. Equation 5 shows that $S^{(n,l)} = C - G^{(n,l)}$. Hence, the distribution of the available spaces after alighting satisfies

$$s_k^{(n,l)} = g_{C-k}^{(n,l)} \quad \forall k = 0, 1, \dots, C. \quad (\text{EC.2})$$

Equations EC.1 and EC.2 hold for every vehicle run. When the steady state limits defined in Section 4.1 exist, taking $l \rightarrow \infty$ converts these run specific relations into their steady state counterparts and thereby establishes Proposition 1.

EC.3.2. Proof of Proposition 2

After alighting and boarding at station n , the steady state vehicle load satisfies $V^{(n)} = \min\{C, G^{(n)} + Q^{(n)}\}$. Define $b_{ij}^{(n)} := \mathbb{P}(V^{(n)} = j \mid G^{(n)} = i)$. If $0 \leq i \leq j < C$, then $b_{ij}^{(n)} = \mathbb{P}(Q^{(n)} = j - i \mid G^{(n)} = i)$. By Assumption 2, this conditional probability equals the marginal probability $q_{j-i}^{(n)}$, which gives the first case in Eq. 15.

If $j = C$ and $0 \leq i < C$, the vehicle reaches capacity when $Q^{(n)} \geq C - i$. Assumption 2 therefore gives $b_{iC}^{(n)} = \mathbb{P}(Q^{(n)} \geq C - i) = 1 - \sum_{k=0}^{C-i-1} q_k^{(n)}$. If $i = j = C$, the vehicle is already full after alighting, so $b_{CC}^{(n)} = 1$. All other transitions are infeasible and have probability zero.

The law of total probability now gives $v_j^{(n)} = \sum_{i=0}^C g_i^{(n)} b_{ij}^{(n)}$ for every $j = 0, \dots, C$. Writing these equations in matrix form yields $v^{(n)} = g^{(n)} B^{(n)}$, which completes the proof.

EC.3.3. Proof of Proposition 3

The proof adapts the transform argument of Powell (1981) to the arbitrary available capacity distribution $s_0^{(n)}, \dots, s_C^{(n)}$, consistent with the variable capacity transform of Powell (1985).

We begin with the one step transition of the queue. From the relationship in Eq. 3,

$$q_k^{(n,l+1)} = \sum_{i=0}^k r_i^{(n,l)} y_{k-i}^{(n,l)}, \quad (\text{EC.3})$$

where $r_k^{(n,l)} := \mathbb{P}(R^{(n,l)} = k)$ and $y_k^{(n,l)} := \mathbb{P}(Y^{(n,l)} = k)$ for every nonnegative integer k . For $u = 0, \dots, C$, also write $r_k^{(n,l)} \Big|_{S^{(n,l)}=u} := \mathbb{P}(R^{(n,l)} = k \mid S^{(n,l)} = u)$. Assumption 3 makes $S^{(n,l)}$ independent of $Q^{(n,l)}$. Conditional on $S^{(n,l)} = u$, all passengers can board when $u \geq Q^{(n,l)}$, so no passenger is left behind. Therefore,

$$r_0^{(n,l)} \Big|_{S^{(n,l)}=u} = \mathbb{P}(u \geq Q^{(n,l)} \mid S^{(n,l)} = u) = \mathbb{P}(u \geq Q^{(n,l)}) = \sum_{k=0}^u q_k^{(n,l)}. \quad (\text{EC.4})$$

If $u < Q^{(n,l)}$, only u passengers can board and $Q^{(n,l)} - u$ passengers remain. Hence,

$$r_i^{(n,l)} \Big|_{S^{(n,l)}=u} = \mathbb{P}(Q^{(n,l)} - u = i \mid S^{(n,l)} = u) = \mathbb{P}(Q^{(n,l)} - u = i) = q_{u+i}^{(n,l)}, \quad i = 1, 2, \dots \quad (\text{EC.5})$$

Using Eqs. EC.4 and EC.5, Eq. EC.3 becomes

$$q_k^{(n,l+1)} = \sum_{u=0}^C s_u^{(n,l)} \left(\sum_{i=0}^u q_i^{(n,l)} y_k^{(n,l)} + \sum_{i=u+1}^{u+k} q_i^{(n,l)} y_{k-i+u}^{(n,l)} \right). \quad (\text{EC.6})$$

Assuming that the steady state probabilities exist, we have $\lim_{l \rightarrow \infty} q_k^{(n,l)} = q_k^{(n)}$, $\lim_{l \rightarrow \infty} s_k^{(n,l)} = s_k^{(n)}$, and $\lim_{l \rightarrow \infty} y_k^{(n,l)} = y_k^{(n)}$. Taking limits on both sides of Eq. EC.6 gives

$$q_k^{(n)} = \sum_{u=0}^C s_u^{(n)} \sum_{i=0}^u q_i^{(n)} y_k^{(n)} + \sum_{u=0}^C s_u^{(n)} \sum_{i=u+1}^{u+k} q_i^{(n)} y_{k-i+u}^{(n)}. \quad (\text{EC.7})$$

Let $Q(z)$, $R(z)$, and $Y(z)$ denote the probability generating functions (PGFs) of the steady state random variables $Q^{(n)}$, $R^{(n)}$, and $Y^{(n)}$, respectively. Suppressing the station superscript, we write

$$Q(z) = \sum_{k=0}^{\infty} q_k^{(n)} z^k, \quad (\text{EC.8})$$

$$R(z) = \sum_{k=0}^{\infty} r_k^{(n)} z^k = \sum_{k=0}^{\infty} z^k \sum_{u=0}^C s_u^{(n)} \cdot r_k^{(n)} \Big|_{S^{(n)}=u}, \quad (\text{EC.9})$$

$$Y(z) = \sum_{k=0}^{\infty} y_k^{(n)} z^k. \quad (\text{EC.10})$$

Substituting Eqs. EC.4 and EC.5 into Eq. EC.9 gives

$$R(z) = \sum_{u=0}^C s_u^{(n)} \sum_{i=0}^u q_i^{(n)} + \sum_{k=1}^{\infty} z^k \sum_{u=0}^C s_u^{(n)} q_{k+u}^{(n)} \quad (\text{EC.11})$$

$$= \sum_{u=0}^C s_u^{(n)} \left[\sum_{i=0}^u q_i^{(n)} + \frac{1}{z^u} Q(z) - \frac{1}{z^u} \sum_{i=0}^u q_i^{(n)} z^i \right]. \quad (\text{EC.12})$$

Taking the steady state limit in Eq. 3 gives $Q^{(n)} = R^{(n)} + Y^{(n)}$. For a fixed vehicle run, the residual queue $R^{(n)}$ is determined by the queue and service outcomes through the current departure. Under Assumption 1, $Y^{(n)}$ is generated by the exogenous Poisson process during the next independently drawn headway and is therefore independent of this history. Consequently, $R^{(n)}$ and $Y^{(n)}$ are independent in the analytical model, and

$$Q(z) = R(z)Y(z). \quad (\text{EC.13})$$

Combining Eqs. EC.12 and EC.13 yields

$$\begin{aligned} Q(z) &= \frac{Y(z) \sum_{u=0}^C s_u^{(n)} \left[\sum_{i=0}^u q_i^{(n)} \left(1 - \frac{z^i}{z^u}\right) \right]}{1 - \sum_{u=0}^C s_u^{(n)} \frac{Y(z)}{z^u}} \\ &= \frac{\sum_{u=0}^C s_u^{(n)} \left[\sum_{i=0}^u q_i^{(n)} (z^C - z^{C-u+i}) \right]}{\frac{z^C}{Y(z)} - \sum_{u=0}^C s_u^{(n)} z^{C-u}}. \end{aligned} \quad (\text{EC.14})$$

EC.3.4. Proof of Proposition 4

Section 4.3.1 identifies the C characteristic roots of the denominator but does not determine the queue probabilities. We now use these roots to factor the numerator of $Q(z)$ and then match its polynomial coefficients to obtain $q_0^{(n)}, \dots, q_{C-1}^{(n)}$.

The numerator of $Q(z)$ is a polynomial of degree at most C . In Eq. 16, the term with $i = u$ vanishes, so the coefficient of z^C in this numerator is $\kappa^{(n)} := \sum_{u=1}^C s_u^{(n)} \sum_{i=0}^{u-1} q_i^{(n)}$. Analyticity requires the numerator to cancel all C characteristic roots according to multiplicity. Therefore, the root z_i^* is repeated in the following product whenever its multiplicity exceeds one, and $Q(z)$ can be written as

$$Q(z) = \frac{\kappa^{(n)} (z-1) \prod_{i=1}^{C-1} (z - z_i^*)}{\frac{z^C}{Y(z)} - \sum_{u=0}^C s_u^{(n)} z^{C-u}}. \quad (\text{EC.15})$$

As $z \rightarrow 1$, both $\text{NUM}(z)$ and $\text{DEN}(z)$ approach zero, while $Q(z) \rightarrow 1$. Applying l'Hôpital's rule to this limit gives the following normalization identity.

$$\lim_{z \rightarrow 1} Q(z) = 1 = \lim_{z \rightarrow 1} \frac{\text{NUM}'(z)}{\text{DEN}'(z)} = \frac{\kappa^{(n)} \prod_{i=1}^{C-1} (1 - z_i^*)}{\sum_{u=0}^C s_u^{(n)} u - Y'(1)} \quad (\text{EC.16})$$

$$\Rightarrow \kappa^{(n)} = \frac{\sum_{u=0}^C s_u^{(n)} u - Y'(1)}{\prod_{i=1}^{C-1} (1 - z_i^*)}. \quad (\text{EC.17})$$

Recall that $\bar{Y}^{(n)} := Y'(1) = \mathbb{E}[Y^{(n)}]$ is the mean number of passenger arrivals within a headway and $\bar{S}^{(n)} := \sum_{u=0}^C s_u^{(n)} u = \mathbb{E}[S^{(n)}]$ is the mean number of available spaces in an arriving vehicle. Substituting Eq. EC.17 into Eq. EC.15 gives

$$Q(z) = \frac{(\bar{S}^{(n)} - \bar{Y}^{(n)})(z-1) \prod_{i=1}^{C-1} \frac{z-z_i^*}{1-z_i^*}}{\frac{z^C}{Y(z)} - \sum_{u=0}^C s_u^{(n)} z^{C-u}}. \quad (\text{EC.18})$$

Equating the numerators in Eqs. EC.18 and 16 gives

$$(\bar{S}^{(n)} - \bar{Y}^{(n)})(z-1) \prod_{i=1}^{C-1} \frac{z-z_i^*}{1-z_i^*} = \sum_{u=0}^C s_u^{(n)} \left[\sum_{i=0}^u q_i^{(n)} (z^C - z^{C-u+i}) \right]. \quad (\text{EC.19})$$

Both sides of Eq. EC.19 are polynomials in z , so their coefficients must be equal. Matching the constant terms on both sides gives the first coefficient relation.

$$(-1)(\bar{S}^{(n)} - \bar{Y}^{(n)}) \prod_{i=1}^{C-1} \frac{z_i^*}{z_i^* - 1} = -q_0^{(n)} s_C^{(n)}. \quad (\text{EC.20})$$

Solving this coefficient relation for the empty queue probability gives

$$q_0^{(n)} = \frac{1}{s_C^{(n)}} (\bar{S}^{(n)} - \bar{Y}^{(n)}) \prod_{i=1}^{C-1} \frac{z_i^*}{z_i^* - 1}. \quad (\text{EC.21})$$

The coefficients of the characteristic equation are real, so every nonreal root and its multiplicity are paired with the corresponding conjugate root. The product in Eq. 18 is therefore real. Under stability and irreducibility, the stationary empty queue probability is positive, which establishes the strict positivity of the product.

As a consistency check, consider fixed capacity with $s_C^{(n)} = 1$ and $\bar{S}^{(n)} = C$. Under this specialization, Eq. EC.19 reduces to the following fixed capacity expression.

$$q_0^{(n)} \Big|_{s_C^{(n)}=1} = (C - \bar{Y}^{(n)}) \prod_{i=1}^{C-1} \frac{z_i^*}{z_i^* - 1}. \quad (\text{EC.22})$$

This expression agrees with Chaudhry, Madill, and Briere (1987). We next derive $q_{1:C-1}^{(n)}$. The numerator of Eq. EC.18 can be rewritten as

$$\begin{aligned} (\bar{S}^{(n)} - \bar{Y}^{(n)})(z-1) \prod_{i=1}^{C-1} \frac{z-z_i^*}{1-z_i^*} &= \frac{1}{s_C^{(n)}} (\bar{S}^{(n)} - \bar{Y}^{(n)}) \prod_{i=1}^{C-1} \frac{z_i^*}{z_i^* - 1} \prod_{i=1}^{C-1} \frac{z_i^* - z}{z_i^*} (z-1) s_C^{(n)} \\ &= s_C^{(n)} q_0^{(n)} (z-1) \prod_{i=1}^{C-1} \left(1 - \frac{z}{z_i^*} \right) \\ &= -s_C^{(n)} q_0^{(n)} \prod_{i=0}^{C-1} \left(1 - \frac{z}{z_i^*} \right). \end{aligned} \quad (\text{EC.23})$$

Define $\prod_{i=0}^{C-1} (1 - z/z_i^*) := \sum_{j=0}^C \eta_j^{(n)} z^j$, where $\eta_j^{(n)}$ is the coefficient of z^j for station n . On the right hand side of Eq. EC.19, the coefficient of z^{C-k} is $-\sum_{u=k}^C s_u^{(n)} q_{u-k}^{(n)}$. In Eq. EC.23, the corresponding coefficient is $-s_C^{(n)} q_0^{(n)} \eta_{C-k}^{(n)}$. Matching these coefficients gives

$$s_C^{(n)} q_0^{(n)} \eta_{C-k}^{(n)} = \sum_{u=k}^C s_u^{(n)} q_{u-k}^{(n)}, \quad k = 1, 2, \dots, C-1. \quad (\text{EC.24})$$

For fixed capacity, $s_C^{(n)} = 1$ and $s_k^{(n)} = 0$ for $0 \leq k < C$. Equation EC.24 then reduces to

$$q_{C-k}^{(n)} = q_0^{(n)} \eta_{C-k}^{(n)}, \quad k = 1, 2, \dots, C-1, \quad \text{if } s_C^{(n)} = 1, \quad (\text{EC.25})$$

which agrees with the result of Chaudhry, Madill, and Briere (1987).

Adding the identity $s_C^{(n)} q_0^{(n)} \eta_0^{(n)} = s_C^{(n)} q_0^{(n)}$, where $\eta_0^{(n)} = 1$, converts Eq. EC.24 into the following matrix representation used to solve the queue probabilities.

$$\tilde{\eta}^{(n)} = q_{0:C-1}^{(n)} \Lambda^{(n)}, \quad (\text{EC.26})$$

where $\tilde{\eta}^{(n)} := [s_C^{(n)} q_0^{(n)} \eta_0^{(n)}, s_C^{(n)} q_0^{(n)} \eta_1^{(n)}, \dots, s_C^{(n)} q_0^{(n)} \eta_{C-1}^{(n)}] \in \mathbb{R}^C$ and

$$\Lambda^{(n)} := \begin{bmatrix} s_C^{(n)} & s_{C-1}^{(n)} & \cdots & s_2^{(n)} & s_1^{(n)} \\ 0 & s_C^{(n)} & \cdots & s_3^{(n)} & s_2^{(n)} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & s_C^{(n)} & s_{C-1}^{(n)} \\ 0 & \cdots & \cdots & 0 & s_C^{(n)} \end{bmatrix} \in \mathbb{R}^{C \times C}. \quad (\text{EC.27})$$

Because $s_C^{(n)} > 0$, the triangular matrix $\Lambda^{(n)}$ is invertible. Therefore,

$$q_{0:C-1}^{(n)} = \tilde{\eta}^{(n)} (\Lambda^{(n)})^{-1}. \quad (\text{EC.28})$$

EC.3.5. Proof of Proposition 5

The derivation adapts the transform argument of Powell (1981) and is consistent with the general bulk service result of Powell (1985). For $X \in \{S, Y\}$, the symbols $\bar{X}^{(n)}$, $\overline{\overline{X}}^{(n)}$, and $\overline{\overline{\overline{X}}}^{(n)}$ denote its mean, second central moment, and third central moment, respectively. Define

$$\mathcal{A}(z) := \frac{(\bar{S}^{(n)} - \bar{Y}^{(n)})(z-1)}{\bar{Y}(z) - \sum_{u=0}^C s_u^{(n)} z^{C-u}}, \quad \mathcal{B}_i(z) := \frac{z - z_i^*}{1 - z_i^*}. \quad (\text{EC.29})$$

Equation EC.18 gives $Q(z) = \mathcal{A}(z) \prod_{i=1}^{C-1} \mathcal{B}_i(z)$. At $z = 1$, $Q(1) = \mathcal{A}(1) = \mathcal{B}_i(1) = 1$ and $\mathcal{B}'_i(1) = 1/(1 - z_i^*)$. Differentiating the product then gives $Q'(1) = \mathcal{A}'(1) + \sum_{i=1}^{C-1} 1/(1 - z_i^*)$.

Let $\mathcal{A}_1(z) := (\bar{S}^{(n)} - \bar{Y}^{(n)})(z - 1)$ and let $\mathcal{A}_2(z)$ denote the denominator of $\mathcal{A}(z)$ in Eq. EC.29. We have $\mathcal{A}_1(1) = \mathcal{A}_2(1) = 0$ and $\mathcal{A}'_1(1) = \mathcal{A}'_2(1) = \bar{S}^{(n)} - \bar{Y}^{(n)} > 0$. A Taylor expansion of $\mathcal{A}_2(z)$ around $z = 1$ gives the following two derivative identities.

$$\mathcal{A}'(1) = -\frac{\mathcal{A}_2''(1)}{2\mathcal{A}_2'(1)}, \quad \mathcal{A}''(1) = \frac{3[\mathcal{A}_2''(1)]^2 - 2\mathcal{A}_2'(1)\mathcal{A}_2'''(1)}{6[\mathcal{A}_2'(1)]^2}. \quad (\text{EC.30})$$

Differentiating $\mathcal{A}_2(z)$ and expressing the resulting raw moments in terms of central moments gives

$$\mathcal{A}'(1) = \frac{\bar{\bar{S}}^{(n)} + \bar{\bar{Y}}^{(n)} + (\bar{S}^{(n)} - \bar{Y}^{(n)})[1 + 2(\bar{S}^{(n)} - C)] - (\bar{S}^{(n)} - \bar{Y}^{(n)})^2}{2(\bar{S}^{(n)} - \bar{Y}^{(n)})}. \quad (\text{EC.31})$$

Substitution into $Q'(1)$ yields the mean in Eq. 22.

For the variance, differentiating $\log Q(z) = \log \mathcal{A}(z) + \sum_{i=1}^{C-1} \log \mathcal{B}_i(z)$ twice and then applying $\mathcal{A}(1) = \mathcal{B}_i(1) = 1$ and $\mathcal{B}''_i(1) = 0$ yields the following identity.

$$\text{Var}[Q^{(n)}] = \mathcal{A}''(1) - \mathcal{A}'(1)^2 + \mathcal{A}'(1) + \sum_{i=1}^{C-1} \left(\mathcal{B}'_i(1) - [\mathcal{B}'_i(1)]^2 \right). \quad (\text{EC.32})$$

Using Eq. EC.30, differentiating $\mathcal{A}_2(z)$ through third order, and then converting the resulting raw moments to central moments yields the following expression.

$$\begin{aligned} & \mathcal{A}''(1) - \mathcal{A}'(1)^2 + \mathcal{A}'(1) \\ &= \frac{1}{12(\bar{S}^{(n)} - \bar{Y}^{(n)})^2} \left[-4(\bar{\bar{S}}^{(n)} - \bar{\bar{Y}}^{(n)})(\bar{S}^{(n)} - \bar{Y}^{(n)}) + 3(\bar{\bar{S}}^{(n)} + \bar{\bar{Y}}^{(n)})^2 \right. \\ & \quad \left. - [6(\bar{\bar{S}}^{(n)} - \bar{\bar{Y}}^{(n)}) - 1](\bar{S}^{(n)} - \bar{Y}^{(n)})^2 - (\bar{S}^{(n)} - \bar{Y}^{(n)})^4 \right]. \end{aligned} \quad (\text{EC.33})$$

Finally, $\mathcal{B}'_i(1) - [\mathcal{B}'_i(1)]^2 = -z_i^*/(1 - z_i^*)^2$. Substituting this identity and Eq. EC.33 into Eq. EC.32 yields the variance in Proposition 5.

EC.3.6. Proof of Proposition 7

Without incidents, vehicle l requires $T^{(n)}$ time units to reach station n . Incident onsets occur only during normal travel and follow a Poisson process with rate γ , so the number of incidents during this travel follows a Poisson distribution with mean $\gamma T^{(n)}$. Independent exponential incident durations with rate θ then yield the compound Poisson representation in Eq. 28.

EC.3.7. Proof of Proposition 8

The raw headway moment generating function (MGF) is

$$\begin{aligned}
 M_{\widehat{H}^{(n,l)}}(t) &= \mathbb{E}[e^{t\widehat{H}^{(n,l)}}] = \mathbb{E}\left[e^{t(\bar{H} + \frac{2\gamma T^{(N)}}{\theta F} + I^{(n,l)} - I^{(n,l-1)})}\right] \\
 &= e^{t(\bar{H} + \frac{2\gamma T^{(N)}}{\theta F})} \mathbb{E}[e^{tI^{(n,l)}}] \mathbb{E}[e^{-tI^{(n,l-1)}}] \\
 &= e^{t(\bar{H} + \frac{2\gamma T^{(N)}}{\theta F})} e^{\gamma T^{(n)}(\frac{\theta}{\theta-t}-1)} e^{\gamma T^{(n)}(\frac{\theta}{\theta+t}-1)} \\
 &= e^{t(\bar{H} + \frac{2\gamma T^{(N)}}{\theta F})} e^{\gamma T^{(n)}(\frac{2t^2}{\theta^2-t^2})}, \quad |t| < \theta.
 \end{aligned} \tag{EC.34}$$

The factorization in Eq. EC.34 follows from the independence of the two copies $I^{(n,l)}$ and $I^{(n,l-1)}$ used to define one generic raw headway. Because this is the common marginal law used for every renewal draw, $M_{\widehat{H}^{(n)}}(t) = M_{\widehat{H}^{(n,l)}}(t)$. Assumption 1 separately imposes independence across successive headway draws in the analytical renewal sequence.

EC.3.8. Proof of Proposition 9

Let $\mu = \bar{H} + \frac{2\gamma T^{(N)}}{\theta F}$ and $\sigma = \frac{2\sqrt{T^{(n)}}\gamma}{\theta}$ denote the mean and standard deviation of $\widehat{H}_{\text{Normal}}^{(n)}$, respectively. Because $H_{\text{Normal}}^{(n)}$ is the positive part of this variable, its MGF is

$$\begin{aligned}
 M_{H_{\text{Normal}}^{(n)}}(t) &= \mathbb{E}[e^{tH_{\text{Normal}}^{(n)}}] = \mathbb{P}[\widehat{H}_{\text{Normal}}^{(n)} \leq 0] \cdot e^0 + \int_0^{+\infty} e^{tz} \cdot \phi_{\widehat{H}_{\text{Normal}}^{(n)}}(z) \cdot dz \\
 &= \Phi\left(\frac{-\mu}{\sigma}\right) + \frac{1}{\sigma\sqrt{2\pi}} \int_0^{+\infty} e^{tz + \frac{(z-\mu)^2}{-2\sigma^2}} dz \\
 &= \Phi\left(\frac{-\mu}{\sigma}\right) + e^{\mu t + \frac{\sigma^2 t^2}{2}} \left[1 - \Phi\left(\frac{-\mu}{\sigma} - \sigma t\right)\right].
 \end{aligned} \tag{EC.35}$$

Completing the square in the exponent gives the final equality in Eq. EC.35. Substituting the definitions of μ and σ into this expression establishes Proposition 9.

EC.3.9. Proof of Proposition 10

For a normal variable with mean μ and standard deviation $\sigma > 0$, Eq. 36 gives the expectation of its positive part. Differentiating this expectation with respect to its two parameters gives

$$\frac{\partial \mathbb{E}[H_{\text{Normal}}^{(n)}]}{\partial \mu} = \Phi\left(\frac{\mu}{\sigma}\right) > 0, \quad \frac{\partial \mathbb{E}[H_{\text{Normal}}^{(n)}]}{\partial \sigma} = \phi\left(\frac{\mu}{\sigma}\right) > 0. \tag{EC.36}$$

The mean $\mu = \bar{H} + 2\gamma T^{(N)} / (\theta F)$ and standard deviation $\sigma = 2\sqrt{T^{(n)}}\gamma / \theta$ are nondecreasing in both γ and $1/\theta$. The mean is strictly increasing in γ , and both parameters are strictly increasing in $1/\theta$ when $\gamma > 0$. Equation EC.36 establishes the stated monotonicity. The case $\gamma = 0$ follows by continuity. Figure EC.1 illustrates these two effects.

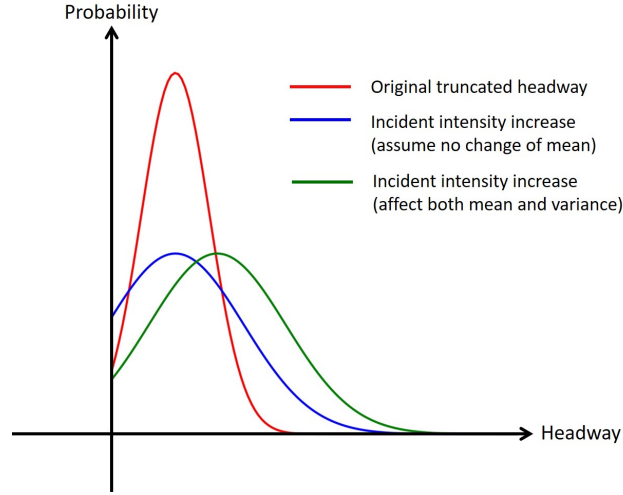


Figure EC.1 Illustration of the impact of incidents on expected headway. As the probability mass at zero does not contribute to the expectation calculation, it is not shown in the figure.

EC.3.10. Proof of Proposition 11

Conditional on $H_{\text{Normal}}^{(n)}$, the variable $Y^{(n)}$ is Poisson with parameter $\lambda^{(n)} H_{\text{Normal}}^{(n)}$. The law of iterated expectations therefore gives $Y(z) = \mathbb{E}\{\exp[\lambda^{(n)} H_{\text{Normal}}^{(n)}(z-1)]\} = M_{H_{\text{Normal}}^{(n)}}[\lambda^{(n)}(z-1)]$. For $\gamma > 0$, substituting $t = \lambda^{(n)}(z-1)$ into the MGF in Proposition 9 gives the stated PGF. When $\gamma = 0$, $H_{\text{Normal}}^{(n)} = \bar{H}$ almost surely, so the conditional Poisson PGF directly gives $Y(z) = \exp[\lambda^{(n)} \bar{H}(z-1)]$.

EC.3.11. Proof of Proposition 12

We establish stability directly from the transient queue dynamics. By the definitions of the queue and available capacity, $R^{(n,l)} = \max\{0, Q^{(n,l)} - S^{(n,l)}\}$. Combining this relation with Eq. 3 gives

$$R^{(n,l+1)} = \max\{0, R^{(n,l)} + Y^{(n,l)} - S^{(n,l+1)}\}. \quad (\text{EC.37})$$

Following Lindley (1952), Eq. EC.37 represents the residual passenger queue after successive vehicle service events as the standard Lindley recursion.

Assumption 1, the independent increments of the Poisson passenger arrival process, and Assumption 3 imply that the net inputs $Y^{(n,l)} - S^{(n,l+1)}$ are identically and independently distributed across vehicle runs. They are exogenous to the current residual queue and have finite mean $\bar{Y}^{(n)} - \bar{S}^{(n)}$. The stability theorem of Loynes (1962) therefore applies directly to Eq. EC.37. If $\bar{Y}^{(n)} < \bar{S}^{(n)}$, the recursion has a proper stationary solution. Because $Q^{(n,l+1)} = R^{(n,l)} + Y^{(n,l)}$, the queue observed at vehicle arrival also has a proper stationary distribution.

If $\bar{Y}^{(n)} > \bar{S}^{(n)}$, the strong law of large numbers gives a positive long run drift for the partial sums of the net inputs, and the residual queue grows without bound. At the boundary $\bar{Y}^{(n)} = \bar{S}^{(n)}$, the

nonconstant independent net inputs in Assumption 3 define a critical Lindley recursion. The critical case has no proper stationary probability distribution under the same result of [Loynes \(1962\)](#). Hence, the station is stable if and only if $\bar{Y}^{(n)} < \bar{S}^{(n)}$. Dividing both sides by $\bar{S}^{(n)} > 0$ and substituting the expressions for the two means gives Eq. 41.

Upstream marginal propagation supplies $s^{(n)}$, and the proof requires only that the resulting station service sequence satisfy Assumption 3. The assumption makes this sequence exogenous to the local queue history. The station by station argument therefore does not presuppose the existence of $q_0^{(n)}$ or another stationary queue probability at the station under consideration.

EC.4. Simulation procedure for comparison

The simulation deliberately relaxes the analytical renewal, marginal route decoupling, and exogenous station service sequence assumptions. For each vehicle, we generate independent compound Poisson exponential incident durations on successive route segments and cumulatively sum them to obtain $I^{(n,l)}$ at each station. Realized departure times follow the recursive FIFO rule in Eq. 9. Vehicle l cannot depart station n before vehicle $l - 1$, so an extreme delay can propagate through several successors. Passenger queues and vehicle loads are updated jointly rather than through independent marginal distributions. When a vehicle arrives, passengers board according to the first come, first served principle up to capacity C . Queue lengths at vehicle arrival and waiting times of boarded passengers are recorded during the simulation.

We evaluate the reference setting and six variations that cover low capacity, long headway, high demand, no incidents, frequent incidents, and long incident duration. Each setting uses 20 independent random seeds. A replication contains 10,000 vehicle runs, and the first 1,000 runs are excluded to reduce initialization effects. We calculate each performance measure separately within a replication and then use variation across replications to construct a 95% confidence interval. The relative ℓ_1 errors compare the mean across replications with the analytical result over stations 1 through 9. Station 10 is excluded from the waiting time evaluation because its zero arrival rate produces no waiting time observations.

Algorithm EC.2 Simulation procedure

```

1: for each of the seven assessment settings do
2:   for each of 20 independent random seeds do
3:     Initialize parameters and empty passenger queues and vehicle loads
4:     for  $l = 1, \dots, 10,000$  do
5:       Set the scheduled dispatch time  $DT^{(l)} \leftarrow (l-1)\overline{H}^{\text{Adj}}$ 
6:       for  $n = 1, \dots, N$  do
7:         Sample the incident duration on segment  $n$  and cumulatively obtain  $I^{(n,l)}$ 
8:          $\hat{t}_D^{(n,l)} \leftarrow DT^{(l)} + T^{(n)} + I^{(n,l)}$ 
9:         if  $l = 1$  then
10:           $t_D^{(n,l)} \leftarrow \hat{t}_D^{(n,l)}$  and initialize the station clock
11:         else
12:           $t_D^{(n,l)} \leftarrow \max\{\hat{t}_D^{(n,l)}, t_D^{(n,l-1)}\}$ 
13:          Sample passenger arrival times during  $(t_D^{(n,l-1)}, t_D^{(n,l)}]$  from the station Poisson process
14:          Record queue length, including passengers left behind by the preceding vehicle
15:          Sample alighting using  $\alpha^{(n)}$  and board passengers by the first come, first served rule up to capacity  $C$ 
16:          Record the residual queue and waiting times of boarded passengers
17:       Exclude the first 1,000 vehicle runs and calculate the four station performance measures
18:   Calculate means and 95% confidence intervals across replications and relative  $\ell_1$  errors over stations 1 through 9

```
