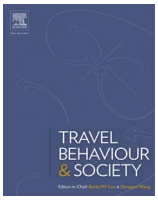
Travel Behaviour and Society 34 (2024) 100689

Contents lists available at ScienceDirect

Travel Behaviour and Society

journal homepage: www.elsevier.com/locate/tbs



Modeling virus transmission risks in commuting with emerging mobility services: A case study of COVID-19

Baichuan Mo [a,*], Peyman Noursalehi [b], Haris N. Koutsopoulos [c], Jinhua Zhao [b]

[a] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States
[b] Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA 20139, United States
[c] Department of Civil and Environmental Engineering, Northeastern University, Boston, MA02115, United States

ARTICLE INFO

Keywords:
COVID-19
Infection risk
Emerging mobility

ABSTRACT

Commuting is an important part of daily life. With the gradual recovery from COVID-19 and more people returning to work from the office, the transmission of COVID-19 during commuting becomes a concern. Recent emerging mobility services (such as ride-hailing and bike-sharing) further deteriorate the infection risks due to shared vehicles or spaces during travel. Hence, it is important to quantify the infection risks in commuting. This paper proposes a probabilistic framework to estimate the risk of infection during an individual's commute considering different travel modes, including public transit, ride-share, bike, and walking. The objective is to evaluate the probability of infection as well as the estimation errors (i.e., uncertainty quantification) given the origin-destination (OD), departure time, and travel mode. We first define a general trip planning function to generate trip trajectories and probabilities of choosing different paths according to the OD, departure time, and travel mode. Then, we consider two channels of infections: 1) infection by close contact and 2) infection by touching surfaces. The infection risks are calculated on a trip segment basis. Different sources of data (such as smart card data, travel surveys, and population data) are used to estimate the potential interactions between the individual and the infectious environment. A first-order approximation is used to simplify the computational complexity. We also derive the closed-form formulation for the estimation errors, enabling us to quantify the uncertainty of the estimation results. The model is implemented in the MIT community as a case study. We evaluate the commute infection risks for employees and students. Results show that most of the individuals have an infection probability close to zero. The maximum infection probability is around 1.3%, implying that the probability of getting infected during the commuting process is low. Individuals with larger travel distances, traveling in transit, and traveling during peak hours are more likely to get infected. Practical implementations of the model are also discussed.

1. Introduction

COVID-19 has greatly affected people's lives all over the world. Recently, with the vaccination and people's prevention consciousness, we are stepping into a new era of living with the virus. With the gradual recovery from the pandemic, more and more people return to work from the office.

Commuting is an important part of the daily lives of people working in an office. In light of the infectiousness of COVID-19, the infection risk during the commuting process is a concern, especially for people using public transportation, as indicated by many previous studies (Mo et al., 2021; Zhou and Koutsopoulos, 2021). On the other hand, recent emerging mobility services (such as ride-hailing and bike-sharing) further deteriorate the infection risks due to shared vehicles or spaces during travel. For example, in ride-hailing, infectious passengers or drivers may be contagious to each other. For bike sharing, infectious riders may contaminate bikes and infect the next users. Therefore, it is important to quantify the infection risks in commuting more broadly, where the results are helpful for people to evaluate their health risks and better inform their commuting route/travel mode choices, and for policymakers to reach informed decisions.

Many previous studies have modeled the COVID-19 infection risks in public transit systems, an important travel mode of commuting. These studies can be categorized from the macro-level at the city scale (Mo et al., 2021) or the micro-scale at the vehicle scale (Shinohara et al., 2021; Zhou and Koutsopoulos, 2021). Researchers have also considered

* Corresponding author.

https://doi.org/10.1016/j.tbs.2023.100689
Received 5 January 2023; Received in revised form 13 August 2023; Accepted 26 September 2023
2214-367X/Published by Elsevier Ltd on behalf of Hong Kong Society for Transportation Studies.

the impact of commuting on the broader spatial transmission of COVID-19 and the related control strategies (Mitze and Kosfeld, 2022; Ando et al., 2021; Fajgelbaum et al., 2021; Kondo, 2021). There are also studies evaluating the impact of COVID-19 on the commuting process with empirical data (such as surveys), such as the impact of COVID-19 on ridership changes, travel mode choices (Tan and Ma, 2021; Medlock et al., 2021), and departure time change (Ecke et al., 2022).

However, there are still two research gaps. First, none of the previous studies have considered the infection modeling for the commuting process as a whole with multiple travel modes and multi-modal trip itineraries. The multi-modal commuting systems are becoming increasingly common in urban areas due to the rise of new mobility services. For example, commuters may use a combination of ride-hailing, bike-sharing, and public transit services to travel to and from work. The unique features of these emerging mobility services, such as the lack of physical barriers between passengers, the varying levels of ventilation, and the limited opportunities for disinfection, may increase the risk of virus transmission during commuting. As such, it is critical to develop infection modeling methodologies that account for these factors. Second, most of the previous studies regarding infection risk modeling only output the probability of infection (or the R_0 value indicating the spreading intensity). The estimation uncertainties (i.e., how accurate are the estimates) are not provided.

In this study, we propose a probabilistic framework to estimate the risk of infection during an individual's commute considering different travel modes, including public transit, ride-share, bike, and walking. The model enables evaluating both the probability of infection and the estimation errors (i.e., uncertainty quantification). We first define a general trip planning function to generate trip trajectories and probabilities of choosing different paths according to the origin, destination, departure time, and travel mode. Two channels of infection are considered: 1) infection by close contact and 2) infection by touching surfaces. The infection risks are calculated on a trip segment basis. Different sources of data (such as smart card data, travel surveys, and population data) are used to estimate the potential interactions between the individual and the infectious environment. A first-order approximation technique is used to simplify the computational complexity. We derive the closed-form formulations for the estimation errors, enabling us to quantify the uncertainties of the estimation results. The model is implemented in the MIT community as a case study. We evaluate the commute infection risks for employees and students. Results show that most of the individuals have an infection probability close to zero. The maximum infection probability is around 1.3%, implying that the probability of getting infected during the commuting process is low. Individuals with larger travel distances, traveling with transit, and traveling during peak hours are more likely to get infected.

The main contribution of this paper is twofold:

- This is the first study dedicated to virus transmission modeling during commuting with the consideration of various travel modes and multi-modal trip itineraries. Infections due to close contact and touching surfaces are both captured. Existing research has largely focused on modeling transmission risks for single travel modes, such as public transit, and only considered close contact infection.
- With multi-modal trip itineraries, this paper proposes a unified framework to model infection risks of both indoor and outdoor travel modes (e.g., walking and bicycling). Outdoor infection is usually overlooked by the literature.
- In addition to estimating infection probabilities, this paper also calculates the estimation errors (i.e., the standard deviation of the estimated probabilities) for uncertainty quantification, which has not been done in the literature.

The remainder of the paper is organized as follows. The literature review is shown in Section 2. In Section 3, we describe the problem and discuss the solution methods. We apply the proposed framework to the MIT community as a case study in Section 4. Finally, we conclude our study and summarize the main findings in Section 5.

2. Literature review

2.1. Infection modeling in public transit

Public transit is an important travel mode for commuting. Previous studies have explored epidemic spreading and infection risk modeling in transit networks. Mo et al. (2021) propose a time-varying weighted encounter network to model the spreading of infectious diseases through public transit systems. The model is implemented at the metropolitan level for population infection calculation. There are many models dedicated to indoor infection modeling in a transit vehicle (Zhou and Koutsopoulos, 2021; Dai and Zhao, 2020; Zhou and Koutsopoulos, 2022; Shen et al., 2021; Das and Ramachandran, 2021; Ku et al., 2021; Shinohara et al., 2021; Zhao et al., 2022). For example, Zhou and Koutsopoulos (2021) propose a modified Wells-Riley model for infection probability calculation in public transportation systems at the vehicle level. The model captures the spatial and temporal passenger flow characteristics in terms of the number of boarding and alighting passengers and the number of infectors. Similarly, Ku et al. (2021) analyzed the degree of infection exposure in public transport by simulating how passengers encounter and infect each other during their journeys. Shinohara et al. (2021) adopted a two-zone-based exponential model to calculate the infection risks in commuter trains by collecting air exchange rate data under various conditions. Zhao et al. (2022) developed a Wells-Riley model-based method to quantitatively evaluate the infection risk of riding public transit. They compared the effectiveness of different countermeasures in managing the spread. In Section 4.5.2, we compare our model results with six studies modeling infection probability in public transit. People can refer to Gartland et al. (2022) for a more comprehensive review of infection modeling in public transit.

2.2. Infection of other diseases in small and confined spaces

Before COVID, there are many studies focusing on estimating the infection risks in small and confined spaces (e.g., transit vehicles). For example, Riley et al. (1978) proposed a model and used it for the epidemiological study of a measles outbreak. The model, known as the Wells-Riley model, has been used extensively by researchers to study the transmission risk of various respiratory diseases, including influenza, SARS, etc. inside aircraft, hospital wards, classrooms, and transit vehicles (Fennelly and Nardell, 1998; Rudnick and Milton, 2003; Ko et al., 2004; Furuya, 2007; Chen et al., 2011; Andrews et al., 2013). Most of these studies focus on an indoor environment, the outdoor infection modeling (such as walking and bicycling) is limited.

2.3. Impact of COVID-19 on commuting

COVID-19 may affect the commuting process in many aspects, such as ridership and service frequency decrease, changes in passengers' travel mode choices, route choices, and departure times choices. Previous studies have evaluated these impacts using different sources of empirical data. For example, many studies have used the smart card data to analyze the impact of COVID-19 on transit ridership changes (Ahangari et al., 2020; Chang et al., 2021; Wilbur et al., 2020; Jenelius and Cebecauer, 2020). There are also studies on ridership changes in ride-hailing systems (Meredith-Karam et al., 2021) and bike-sharing systems (Wang and Noland, 2021). Tan and Ma (2021) conducted a survey to understand commuters' mode choice changes during the COVID-19 pandemic. They used a logistic regression model with personal attributes, travel attributes, and perception of COVID-19 based on a sample of 559 responses to a survey. Ecke et al. (2022) examine how people's commuting behavior changed before and after COVID-19. The results show that people did not significantly change their commuting

behavior in terms of commuting time and commuting mode.

2.4. Impact of commuting on COVID-19 spreading

Commuting may contribute to the spreading of COVID-19 by transporting infectious passengers across different regions. For example, Mitze and Kosfeld (2022) proposed a spatial econometric model of the epidemic spread to identify the role played by commuting-to-work patterns for spatial disease transmission and explored if the imposed containment policies affected the strength of this transmission channel. Ando et al. (2021) investigated the relationship among commuting, the risk of COVID-19, and COVID-19-induced anxiety using internet-based survey data from 27,036 respondents. Fajgelbaum et al. (2021) designed an optimal dynamic lockdown strategy against COVID-19 within a commuting network. Kondo (2021) developed a spatial susceptible-exposed-infectious-recovered model to analyze the effects of restricting inter-regional commuting mobility on the spatial spread of the COVID-19 infection in Japan. Zhu et al. (2022) considered the effects of different travel modes and travel destinations on COVID-19 transmission in global cities.

2.5. Research gaps

To date, no studies have comprehensively modeled virus transmission risks during commuting processes that involve multiple travel modes and complex itineraries. Existing research has largely focused on modeling transmission risks for single travel modes, such as public transit, or on epidemic spreading at the city level, without using travel mode- or itinerary-specific methodologies. Moreover, previous studies have not accounted for the unique challenges of modeling transmission risks in multi-modal systems, which include outdoor travel modes like walking and bicycling. Furthermore, while many studies have estimated the probability of infection or the R0 value as a measure of transmission intensity, few have reported on the accuracy of these estimates or the uncertainties associated with their modeling methodologies.

3. Methodology

3.1. Problem definition

Consider a set of individuals $\mathscr{I}$. For each individual $i \in \mathscr{I}$, suppose that we know their origin o_i, destination d_i, departure time t_i, and travel mode m_i for their daily commuting. The objective of this study is to estimate the probability of individual i getting affected: $\mathbb{P}(i \text{ infected}|o_i, d_i, t_i, m_i)$. Besides, we also aim to quantify the estimation uncertainty. If we treat $\mathbb{P}(i \text{ infected}|o_i, d_i, t_i, m_i)$ as a random variable, another goal of the study is to obtain the standard error of the estimation: $\sqrt{\text{Var}[\mathbb{P}(i \text{ infected}|o_i, d_i, t_i, m_i)]}$.

In the following sections, we first illustrate how $\mathbb{P}(i \text{ infected}|o_i, d_i, t_i, m_i)$ is estimated. The estimation of standard errors is illustrated in Section 3.5.

3.2. Trip itinerary generation

Given an individual i's trip information (o_i, d_i, t_i, m_i), there exists a trip planner function $TP(\cdot)$ that takes (o_i, d_i, t_i, m_i) as input, and outputs a set of feasible paths for the individual $\mathscr{R}_i$ and the associated path choice probability $\mathbb{P}_{\text{Path}}(r)$ for all paths $r \in \mathscr{R}_i$, that is:

$$\mathscr{R}_i,\ (\mathbb{P}_{\text{Path}}(r))_{r\in\mathscr{R}_i} = TP(o_i, d_i, t_i, m_i) \tag{1}$$

For example, we can define $TP(\cdot)$ as a composed function of the Google Map API and a C-logit model:

$$TP(o_i, d_i, t_i, m_i) = \texttt{C-Logit}\circ\texttt{GoogleMapAPI}(o_i, d_i, t_i, m_i) \tag{2}$$

Specifically, the Google Map API returns the path set $\mathscr{R}_i$ and path attributes $\boldsymbol{X}_r$ for all $r \in \mathscr{R}_i$ (e.g., travel time, travel cost, etc.), that is:

$$\mathscr{R}_i,\ (\boldsymbol{X}_r)_{r\in\mathscr{R}_i} = \texttt{GoogleMapAPI}(o_i, d_i, t_i, m_i) \tag{3}$$

The C-Logit model outputs the path choice probability and the standard errors for each path. The C-Logit model is an extension of the multinomial logit (MNL) model to correct for the correlation among paths due to overlapping (Cascetta et al., 1996; Mo et al., 2023). The key idea is to define a term called the "commonality factor" of path r (i.e., CF_r), which measures the degree of similarity of path r with the other paths of the same OD. Based on the C-logit model, the probability of choosing path r can be calculated as

$$\mathbb{P}_{\text{Path}}(r) = \texttt{C-Logit}(\boldsymbol{X}_r) = \frac{\exp[\boldsymbol{\beta}^T\cdot(\boldsymbol{X}_r, CF_r)]}{\sum_{r'\in\mathscr{R}_i}\exp[\boldsymbol{\beta}^T\cdot(\boldsymbol{X}_{r'}, CF_{r'})]} \tag{4}$$

where $\boldsymbol{\beta}$ is the parameter vector to estimate. The formulation of CF_r can be found in Cascetta et al. (1996).

Note the $TP(\cdot)$ can be defined more broadly than Google Map API plus the C-logit model. Other examples include the k-shortest path in a multiple-modal network compounded with any behavioral model for path choices (Mo et al., 2022).

3.3. Infection modeling for a trip segment

3.3.1. Definition of a trip segment

A path $r \in \mathscr{R}_i$ usually contains multiple trip segments, such as walking from home to a bus station, taking a bus, and walking from a bus station to the office. The infection may happen at every trip segment. In this study, we define a segment s of a path as continuous travel with the same travel mode along the path. Let the set of all segments for path r be $\mathscr{S}_r$. It is worth noting that, if a transit trip has one or more transfers, we separate the transit trip into multiple segments based on transfers because the passenger needs to first alight and then board a new vehicle, which is equivalent to changing to a "new travel mode" in infection modeling. We also ignore short walking segments (less than 3 min or less than 1 km, e.g., transfer walking) for modeling convenience.

We provide three examples to illustrate the definition of segments (Fig. 1). The first example shows three segments: walking from home to a subway station, taking the subway, and walking from a subway station to the office. The second example is a ride-hailing trip with only a single segment. The third example shows a transit trip with a transfer, which is separated into two segments by definition.

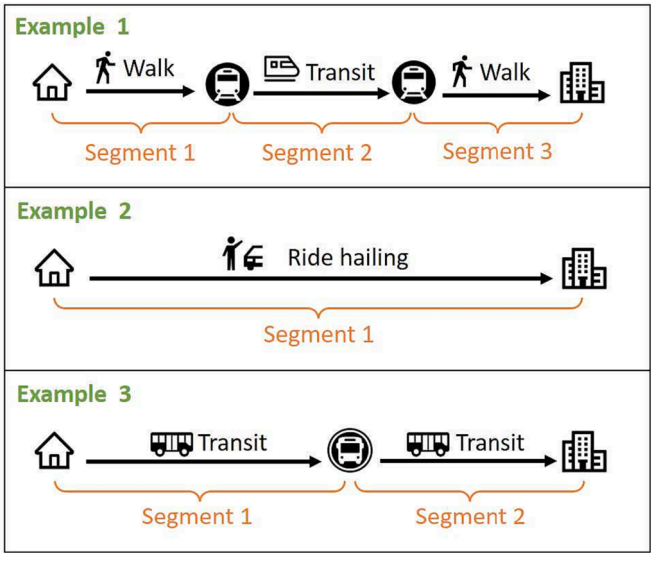


Fig. 1. Illustration of the segment definition.

3.3.2. Infection risk for each segment

We consider two different channels for infection: 1) infected by close contact with infectious persons and 2) infected by touching infectious surfaces.

Infection by close contact: Consider a trip segment $s \in \mathscr{S}_r$. We define $\mathscr{P}_s$ as the set of persons that have been in the six feet infectious range of individual i. For a person $p \in \mathscr{P}_s$, let the duration of the interaction between i and p be $d_{i,p}$. If p is infectious, i would have a probability of getting infected. Depending on whether the interaction happens indoors (e.g., in a bus) or outdoors (e.g., walking by), there are two different physical models.

In an indoor environment, the probability of i getting infected by p can be calculated by the well-known Wells-Riley model (Riley et al., 1978; Fennelly and Nardell, 1998):

$$\mathbb{P}_{\text{Indoor}}\left(p \to i \middle| p \text{ infectious}\right) = 1 - \exp\left(\frac{-b \cdot q \cdot (1-\alpha) \cdot d_{i,p}}{Q^{\text{Indoor}}}\right) \tag{5}$$

where b is the breathing rate per person (m^3/hour); q is the quanta generation rate (/hour), $\alpha \in [0,1]$ is the effectiveness of mask or face covering, Q is the room ventilation rate of clean air (m^3/hour).

In an outdoor environment, Rowe et al. (2021) proposed an airshed model that derives a similar infection probability formulation:

$$\mathbb{P}_{\text{Outdoor}}\left(p \to i \middle| p \text{ infectious}\right) = 1 - \exp\left(\frac{-b \cdot q \cdot (1-\alpha) \cdot d_{i,p}}{Q^{\text{Outdoor}}}\right) \tag{6}$$

where $Q^{\text{Outdoor}} = L \cdot H \cdot V_\infty$ is the outdoor ventilation rate of clean air (as shown in Fig. 2). H and L are the height and length (perpendicular to the wind direction) for a hypothetical outdoor modeling space. V_∞ is the wind velocity (m/h). The suggested values for H and L are around $5m$ and $50m$, respectively.

Hence, given the set of contact persons, the total infectious probability of i during the segment s due to close contact is

$$\begin{aligned}\mathbb{P}^{(s)}_{\text{Cont}}\left(i \text{ infected}\right) = 1 - \prod_{p \in \mathscr{P}_s} \big[&\left(1 - \mathbb{P}_{\text{In/Outdoor}}\left(p \to i \mid p \text{ infectious}\right)\right) \\ &\cdot \mathbb{P}\left(p \text{ infectious}\right) + \mathbb{P}\left(p \text{ not infectious}\right) \big]\end{aligned} \tag{7}$$

where Eq. 7 is due to the fact that the probability of getting infected by at least one of $p \in \mathscr{P}_s$ equals one minus the probability of not getting infected by anyone.

Infection by touching surfaces: Wilson et al. (2021) estimate the infection probability of a single hand-to-fomite touch as a function of viral bioburden. Their estimation already accounts for uncertainties in transfer efficiency, fractions of the hand used for surface and face contact, and surface areas of the hand and of fomites available for contact. Define this probability as $\mathbb{P}_{\text{Touch}}(i \text{ infected} | V)$, where V is the viral bioburden of this touch. For simplicity, let us assume the viral bioburden during a trip segment is a constant and the value is V_s. We also assume that the number of touches during segment s (defined as N_s) is proportional to the duration of travel in s (defined as T_s), and the factor is γ_s (i.e., $N_s = \gamma_s \cdot T_s$). On another hand, people who wear masks are less likely to generate viral bioburden. The effective viral bioburden can be expressed as $V_s \cdot (1-\alpha)$. Therefore, the total infection probability for individual i due to surface touching is:

$$\mathbb{P}^{(s)}_{\text{Surf}}\left(i \text{ infected}\right) = 1 - \left(1 - \mathbb{P}_{\text{Touch}}(i \text{ infected} \mid V_s \cdot (1-\alpha))\right)^{N_s} \tag{8}$$

The empirical values of V_s can be obtained from Harvey et al. (2020) who collected viral bioburden data in daily activity environments. Since the infection risk for a single touch is small, Eq. 8 can be approximated by first-order Taylor series:

$$\mathbb{P}^{(s)}_{\text{Surf}}\left(i \text{ infected}\right) \approx \gamma_s \cdot T_s \cdot \mathbb{P}_{\text{Touch}}\left(i \text{ infected} \,\middle|\, V_s \cdot \left(1-\alpha\right)\right) \tag{9}$$

where Eq. 9 is computationally more efficient than Eq. 8.

3.3.3. Infection risk by different travel modes

Each segment $s \in \mathscr{S}_r$ is associated with a specific travel mode. Though we provide a general infection risk calculation model in Section 3.3.2, it is important to specify the model parameters and variables for each travel mode. In this section, we assume that the infectious environment is the same during the travel in a segment s and the values are $b_s, q_s, Q_s^{\text{Indoor}}$, and Q_s^{Outdoor}.

Infection risk of transit segment: A transit segment (either bus or rail) usually includes multiple stops. Let the set of stops for the transit segment s except for the last one being $\mathscr{A}_s$. We exclude the last stop because individual i will alight when he/she arrives at the last stop. Let the set of passengers in a vehicle (exclude individual i) when the vehicle departs from station $a \in \mathscr{A}_s$ be $\mathscr{P}_{s,a}$. $\mathscr{P}_{s,a}$ can be obtained by smart card data and $\mathscr{P}_s = \cup_{a \in \mathscr{A}_s} \mathscr{P}_{s,a}$. Let the vehicle travel time from station a to the next stop be TT_a.

Then the infection probability for individual i due to close contact in the vehicle when it travels from station a to the next stop is:

$$\begin{aligned}\mathbb{P}^{(s,a)}_{\text{Cont}}\left(i \text{ infected}\right) = 1 - \prod_{p \in \mathscr{P}_{s,a}} \Bigg[&\mathbb{P}\left(p \text{ infectious}\right) \cdot \exp\left(\frac{-b_s \cdot q_s \cdot (1-\alpha) \cdot TT_a}{Q_s^{\text{Indoor}}}\right) \\ &+ \left(1 - \mathbb{P}\left(p \text{ infectious}\right)\right) \Bigg]\end{aligned} \tag{10}$$

For any $p \in \mathscr{P}_{s,a}$, we can use smart card data to obtain their origin stations. Hence, $\mathbb{P}(p \text{ infectious})$ and $\mathbb{P}(p \text{ not infectious})$ can be approximated by regional infection statistics based on their origin stations. The total infection probability in the transit segment is:

$$\mathbb{P}^{(s)}_{\text{Cont}}\left(i \text{ infected}\right) = 1 - \prod_{a \in \mathscr{A}_s} \left(1 - \mathbb{P}^{(s,a)}_{\text{Cont}}\right) \quad \text{when } s \text{ is a transit segment} \tag{11}$$

Eqs. 10 and 11 may be computationally inefficient. When $\mathbb{P}(p\text{infectious})$ is small, we can use the following approximation by ignoring all second-order multiplication terms with $(\mathbb{P}(p\text{infectious}))^2$:

$$\begin{aligned}&\prod_{p \in \mathscr{P}_{s,a}} \left[\mathbb{P}\left(p\text{infectious}\right) \cdot \exp\left(\frac{-b_s \cdot q_s \cdot (1-\alpha) \cdot TT_a}{Q_s^{\text{Indoor}}}\right) + \left(1 - \mathbb{P}\left(p\text{infectious}\right)\right)\right] \\ = &\prod_{p \in \mathscr{P}_{s,a}} \left[1 - \mathbb{P}\left(p\text{infectious}\right) \cdot \left(1 - \exp\left(\frac{-b_s \cdot q_s \cdot (1-\alpha) \cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right)\right] \\ \approx &1 - \sum_{p \in P_{s,a}} \mathbb{P}\left(p\text{infectious}\right) \cdot \left(1 - \exp\left(\frac{-b_s \cdot q_s \cdot (1-\alpha) \cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right)\end{aligned} \tag{12}$$

Then we have

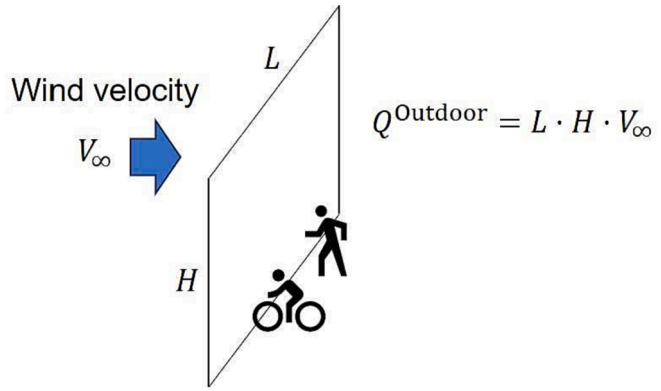


Fig. 2. Illustration of the outdoor ventilation rate.

$$\mathbb{P}^{(s,a)}_{\text{Cont}}\left(i\,\text{infected}\right) \approx \sum_{p\in\mathscr{P}_{s,a}} \mathbb{P}\left(p\,\text{infectious}\right)\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right) \tag{13}$$

which is simply the summation of probabilities of getting infected by anyone in $\mathscr{P}_{s,a}$. Similarly, if $\mathbb{P}^{(s,a)}_{\text{Cont}}$is small, we can approximate $\mathbb{P}^{(s)}_{\text{Cont}}$ as:

$$\mathbb{P}^{(s)}_{\text{Cont}}\left(i\text{infected}\right) \approx \sum_{a\in\mathscr{A}_s}\sum_{p\in\mathscr{P}_{s,a}} \mathbb{P}\left(p\text{infectious}\right)\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right)$$
$$\text{when}\,s\,\text{is a transit segment} \tag{14}$$

where Eqs. 13 and 14 are computationally more efficient because we replace the production with a summation.

In terms of the surface-touching infection, we only need to specify V_s (viral bioburden), γ_s (touching rate), and T_s (travel time) to use Eq. 9. It is worth noting that V_s can vary across different times of the day and transit routes according to the demand level. In general, times and routes with higher demand should have higher V_s.

Infection risk of walk/bike segment: For a walk or a personal bike (not bike-sharing) segment, we assume there is no surface-touching infection risk because the commuter does not need to touch public surfaces during commuting:

$$\mathbb{P}^{(s)}_{\text{Surf}}\left(i\text{infected}\right)=0\ \text{when}\,s\,\text{is a personal bike}\left(\text{not bike sharing}\right)\text{or walk segment} \tag{15}$$

For a bike-sharing segment, the previous riders may be infectious and thus the bike is contaminated. We can specify V_s and γ_s and then use Eq. 9 for the infection risk calculation. However, it is worth noting that V_s for bike-sharing should be much smaller than that of transit.

For the close-contact infection, we can approximate $\mathscr{P}_s$ from the pedestrian density data set. Denote the average contact time for an encounter as $d_{\text{W/B}}$. With the same approximation techniques, we can calculate the infection probability as

$$\mathbb{P}^{(s)}_{\text{Cont}}\left(i\,\text{infected}\right) \approx |\mathscr{P}_s|\cdot\mathbb{P}\left(p\,\text{infectious}\right)\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot d_{\text{W/B}}}{Q_s^{\text{Outdoor}}}\right)\right)$$
$$\text{when}\,s\,\text{is a bike/walk segment} \tag{16}$$

where $\mathbb{P}(p\ \text{infectious})$ can be obtained from the regional infectious statistics based on the walk/bike routes; $|\mathscr{P}_s|$ is the average number of encounters during the segment.

Infection risk of ride-hailing segment: When commuting with ride-hailing, $\mathscr{P}_s$ includes the driver and potential shared riders. $\mathscr{P}_s$ can be obtained from ride-hailing open source data to approximate the average number of persons in a vehicle. The contact time $d_{i,p}$ for the driver and individual i is T_s (total segment travel time). While $d_{i,p}$ for individual i and shared riders will be approximated by shared rides data. With the same first-order approximation, we can calculate the infection probability due to close contact as

$$\mathbb{P}^{(s)}_{\text{Cont}}\left(i\,\text{infected}\right) \approx \sum_{s\in\mathscr{P}_s} \mathbb{P}\left(p\,\text{infectious}\right)\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot d_{i,p}}{Q_s^{\text{Indoor}}}\right)\right)$$
$$\text{when}\ s\ \text{is a ride hailing segment} \tag{17}$$

In terms of surface-touching risks, similar to bike-sharing, it is possible that the driver or previous riders are infectious and thus the seats are contaminated. Eq. 9 can be used after specifying the corresponding parameters. V_s for ride-hailing should be smaller than that of transit but greater than that of bike-sharing.

Infection risk of driving segment: Driving is a private commute mode and there is generally no close contact or infectious surface touching during the travel. Hence, we simply assume $\mathbb{P}^{(s)}_{\text{Surf}}(i\ \text{infected}) = \mathbb{P}^{(s)}_{\text{Cont}}(i\ \text{infected}) = 0$ when s is a driving segment.

3.4. Infection risk integration

Section 3.3 provides the formulations for the calculation of infection risks for a specific segment. Since a commuting path r consists of multiple segments, the total infection probability if using path $r \in \mathscr{R}_i$ is:

$$\begin{aligned}\mathbb{P}(i\ \text{infected}|r) &= 1-\prod_{s\in\mathscr{I}_r}\mathbb{P}\left(i\ \text{not infected in segment}\ s\right)\\ &= 1-\prod_{s\in\mathscr{I}_r}\left(1-\mathbb{P}^{(s)}_{\text{Surf}}\left(i\ \text{infected}\right)\right)\cdot\left(1-\mathbb{P}^{(s)}_{\text{Cont}}\left(i\ \text{infected}\right)\right)\end{aligned} \tag{18}$$

Similarly, if the infectious probability is small in each segment, we have:

$$\mathbb{P}\left(i\ \text{infected}\middle| r\right) \approx \sum_{s\in\mathscr{I}_r}\left(\mathbb{P}^{(s)}_{\text{Cont}}\left(i\ \text{infected}\right)+\mathbb{P}^{(s)}_{\text{Surf}}\left(i\ \text{infected}\right)\right) \tag{19}$$

Combining with the path choice probabilities from Section 3.2, we get the total infectious probability of individual i during commuting:

$$\mathbb{P}\left(i\ \text{infected}\middle| o_i,d_i,t_i,m_i\right) = \sum_{r\in\mathscr{R}_i}\mathbb{P}_{\text{Path}}\left(r\right)\cdot\mathbb{P}\left(i\ \text{infected}\middle| r\right) \tag{20}$$

3.5. Uncertainty quantification

Though Eq. 20 provides the final infection probability for a given trip, the estimation may suffer from errors due to uncertainties in parameters. In this section, we treat $\mathbb{P}(i\ \text{infected} \mid o_i,d_i,t_i,m_i)$ as a random variable and aim to calculate its standard errors: $\sqrt{\text{Var}[\mathbb{P}(i\ \text{infected} \mid o_i,d_i,t_i,m_i)]}$.

In this study, we focus on the uncertainty due to the infection calculations. Hence, let us assume $\mathbb{P}_{\text{Path}}(r)$ is fixed. Then:

$$\text{Var}\left[\mathbb{P}\left(i\ \text{infected}\ \middle|\ o_i,d_i,t_i,m_i\right)\right] = \sum_{r\in\mathscr{R}_i}(\mathbb{P}_{\text{Path}}(r))^2\cdot\text{Var}\left[\mathbb{P}\left(i\ \text{infected}\ \middle|\ r\right)\right] \tag{21}$$

From Eq. 19, we can further decompose the variance to different segments due to the independence across segments:

$$\begin{aligned}\text{Var}\left[\mathbb{P}\left(i\ \text{infected}\ \middle|\ r\right)\right] = \sum_{s\in\mathscr{I}_r}\Big(&\text{Var}\left[\mathbb{P}^{(s)}_{\text{Cont}}\left(i\ \text{infected}\right)\right]\\ &+\text{Var}\left[\mathbb{P}^{(s)}_{\text{Surf}}\left(i\ \text{infected}\right)\right]\Big)\end{aligned} \tag{22}$$

Therefore, we only need to specify $\text{Var}\left[\mathbb{P}^{(s)}_{\text{Cont}}(i\ \text{infected})\right]$ and $\text{Var}\left[\mathbb{P}^{(s)}_{\text{Surf}}(i\ \text{infected})\right]$ for different types (i.e., travel modes) of segments.

Transit segment: According to Eq. 14, the infection probability due to close contact in public transit is simply the probability summation over different stations and encounters. In the real world, $\mathscr{P}_{s,a}$ is uncertain due to demand variations in public transit systems. However, it is generally hard to model the uncertainty of a set. Therefore, let us define the average probability that a passenger in $\mathscr{P}_{s,a}$ is infectious as $\overline{\mu}_{s,a}$. Then Eq. 14 can be rewrite as:

$$\mathbb{P}^{(s)}_{\text{Cont}}\left(i \text{ infected}\right) \approx \sum_{a\in\mathscr{A}_s}\left|\mathscr{P}_{s,a}\right|\cdot\overline{\mu}_{s,a}\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right)$$
when s is a transit segment (23)

where $|\mathscr{P}_{s,a}|$ is the number of passengers in the vehicle (excluding i) when the vehicle departs from station a. The variance can be formulated as:

$$\text{Var}\left[\mathbb{P}^{(s)}_{\text{Cont}}\left(i \text{ infected}\right)\right] = \sum_{a\in\mathscr{A}_s}\text{Var}\left[\left|\mathscr{P}_{s,a}\right|\cdot\overline{\mu}_{s,a}\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right)\right]$$
when s is a transit segment (24)

Deriving the closed-form formulation for the variance of the product of several random variables (or the exponential of several variables) is difficult. In some simple cases, one may use Jensen's inequality to get the variance lower (or upper) bound. However, consider a general function $f(Z)$ (may not be convex or concave), where Z is a vector of random variables. Obtaining the analytical form of $\text{Var}[f(Z)]$ is generally hard. Therefore, we propose a bootstrapping-based algorithm to estimate the empirical variance (Algorithm 1). The inputs of the algorithm are $f(Z)$, the distribution of Z (i.e., $\mathbb{P}(Z)$), and maximum sample times M. The idea is to sample $Z \sim \mathbb{P}(Z)$ and use samples to estimate the variance.

Algorithm 1 Bootstrapping-based empirical variance estimation algorithm

1: **function** Bootstrapping-Variance$f(Z)$, $\mathbb{P}(Z)$, M
2: **for** $m = 1, 2, \ldots, M$
3: Sample $Z^{(m)} \sim \mathbb{P}(Z)$
4: $\bar{f}(Z) = \frac{1}{M}\sum_{m=1}^{M} f(Z^{(m)})$ ▷Estimate empirical mean
5: $\text{Var}[f(Z)] = \frac{1}{M}\sum_{m=1}^{M}\left(f(Z^{(m)}) - \bar{f}(Z)\right)^2$ ▷Estimate empirical variance
6: **return** $\text{Var}[f(Z)]$

It is worth noting that, the sampling process in Algorithm 1 (Line 3) can be done independently, jointly, or sequentially, depending on whether the elements in Z are independent, dependent, or conditionally independent.

Given Algorithm 1, we can estimate $\text{Var}\left[\left|\mathscr{P}_{s,a}\right|\cdot\overline{\mu}_{s,a}\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot TT_a}{Q_s^{\text{Indoor}}}\right)\right)\right]$ through the distribution of $|\mathscr{P}_{s,a}|, \overline{\mu}_{s,a}, b_s, q_s, TT_a$, and Q_s. In this study, we assume that $|\mathscr{P}_{s,a}|$ (vehicle load) and TT_a (travel time) are normally distributed based on the observations in the empirical data. Their distribution parameters can be estimated from the smart card and automated vehicle location data. b_s, q_s, and Q_s are uniformly distributed and their distributions are shown in Table 1 based on the literature. For $\overline{\mu}_{s,a}$, the distribution is hard to parameterize because it is generated by sampling different $\mathscr{P}_{s,a}$. We, therefore, keep all samples for $\overline{\mu}_{s,a}$ as an empirical distribution. Then every sample of $\overline{\mu}_{s,a} \sim \mathbb{P}(\overline{\mu}_{s,a})$ is essentially a bootstrap from its sample pool.

In terms of surface touching infection, from Eq. 9, we use Algorithm 1 to estimate the variance by setting $f(Z) = \gamma_s\cdot T_s\cdot\mathbb{P}_{\text{Touch}}(i \text{ infected}|V_s)$ and $Z = (\gamma_s, T_s, V_s, \mathbb{P}_{\text{Touch}}(i \text{ infected}|V_s))$. The distribution of $\mathbb{P}_{\text{Touch}}(i \text{ infected}|V_s)$ can be obtained from the data in Wilson et al. (2021). The distribution of V_s can be obtained from the results in Harvey et al. (2020). The distribution of T_s can be obtained from the AVL data. $\gamma_s$is assumed to be uniformly distributed and its distribution is given in Table 1.

Walk/Bike segment: The variance of contact-based infection risks in the walk and bike segments, according to Eq. 16, can be expressed as:

$$\text{Var}\left[\mathbb{P}^{(s)}_{\text{Cont}}\left(i \text{ infected}\right)\right] = \text{Var}\left[\left|\mathscr{P}_s\right|\cdot\overline{\mu}_s\cdot\left(1-\exp\left(\frac{-b_s\cdot q_s\cdot(1-\alpha)\cdot d_{\text{W/B}}}{Q_s^{\text{Outdoor}}}\right)\right)\right]$$
when s is a bike/walk segment (25)

where $\overline{\mu}_s$ is the average infectious probability of encounters in $\mathscr{P}_s$. It has a similar formulation as Eq. 24, thus can be calculated using Algorithm 1. The distribution of $|\mathscr{P}_s|$ can be obtained from pedestrian flow data. The distribution of $\overline{\mu}_s$ can be obtained from regional statistics. All infection-related parameters (including $d_{\text{W/B}}$) are assumed to be uniformly distributed and their distributions are shown in Table 1. Since we assume there are no infection risks related to surface touching for the walk and bike segments, we do not need to consider their variances.

Ride hailing segment: For the ride-hailing segment, since $\mathscr{P}_s$ is relatively small (maximum three riders in a vehicle), we assume $|\mathscr{P}_s|$ follows a multinomial distribution across $\{0, 1, 2, 3\}$ (i.e., maximum 2 other passengers plus 1 driver). The specific distribution can be obtained from ride-sharing data. To get $\text{Var}[\mathbb{P}^{(s)}_{\text{Cont}}(i \text{ infected})]$ for ride-hailing segments, the sampling process is as follows (slightly different from Algorithm 1):

- Step 1: Sample $|\mathscr{P}_s|$ from the multinomial distribution.

Table 1
Infection-related parameters.

Mode	Parameters and distribution							
	b_s (m^3/h)	q_s ($/h$)	V_∞ (m/s)	L (m)	H (m)	Q_s^{Indoor} (m^3/h)	γ_s ($/h$)	V_s (gc/cm^2)
Transit (Train)	$\mathscr{U}(0.65, 0.79)$	$\mathscr{U}(100, 300)$	N.A.	N.A.	N.A.	$\mathscr{U}(800, 1100)$	$\mathscr{U}(3, 5)$	$\mathscr{U}(30, 100)$
Transit (Bus)	$\mathscr{U}(0.65, 0.79)$	$\mathscr{U}(100, 300)$	N.A.	N.A.	N.A.	$\mathscr{U}(300, 500)$	$\mathscr{U}(3, 5)$	$\mathscr{U}(150, 300)$
Bike	$\mathscr{U}(1.4, 1.8)$	$\mathscr{U}(100, 300)$	$\mathscr{U}(2, 4)$	$\mathscr{U}(30, 60)$	$\mathscr{U}(2.5, 5)$	N.A.	N.A.	N.A.
Walk	$\mathscr{U}(1.2, 1.6)$	$\mathscr{U}(100, 300)$	$\mathscr{U}(1, 2)$	$\mathscr{U}(30, 60)$	$\mathscr{U}(2.5, 5)$	N.A.	N.A.	N.A.
Ride-hailing	$\mathscr{U}(0.65, 0.79)$	$\mathscr{U}(100, 300)$	N.A.	N.A.	N.A.	$\mathscr{U}(90, 120)$	$\mathscr{U}(1, 3)$	$\mathscr{U}(2, 50)$

- N.A.: Not applicable.
- The parameters for transit infections are obtained from Zhou and Koutsopoulos (2021).
- The parameters for quanta generation rates are similar to Zhou and Koutsopoulos (2021). They are back-calculated using the method mentioned in Stephens (2012) and Rudnick and Milton, 2003. The values are consistent with the findings from studies (Buonanno et al., 2020; Buonanno et al., 2020; Miller et al., 2021), as over 100 quanta/hour.
- The parameters for outdoor infection models are obtained from Rowe et al. (2021).
- The ride-hailing ventilation rate is based on Ott et al. (2008).
- The viral bioburden data is adapted from Harvey et al. (2020) (assuming 10% viruses are infectious).

- Step 2: Based on the value of $|\mathscr{P}_s|$, sample the same number of passengers and drivers to generate the $|\mathscr{P}_s$, for each passenger $p \in \mathscr{P}_s$, we also sample $d_{i,p}$ (shared trip time).
- Step 3: Sample other infection-related parameters from the distribution defined in Table 1. Calculate the infection probability close contact based on Eq. 17.

After M samples, we can estimate the sample variance as an approximation of $\mathrm{Var}[\mathbb{P}^{(s)}_{\mathrm{Cont}}(i \text{ infected})]$ similar to Algorithm 1. The variance of surface touching-based infection probability is calculated in the same way as a transit segment except for replacing the distributions of T_s, V_s, and γ_s.

For the driving segment, since we assume there is no infection risk, the variances are not calculated.

4. Case study

The case study is based on available data from MIT. The model is implemented for the commuting of all MIT students and staff in the greater Boston area.

4.1. Data sources and parameter settings

We use data from various sources in this study to estimate the infection risk and set up parameters. The first data set is the MIT staff commuting survey. The survey collects (o_i, d_i, t_i, m_i) of every individual $i \in \mathscr{I}$. Given this information, for every individual i, we generate the set of paths $\mathscr{R}_i$ and the associated path attributes (i.e., walking time, waiting time, in-vehicle time, travel cost) based on the Google map API. For simplicity, we only consider the first path associated with the travel mode m_i recommended by Google Map API (i.e., ignoring the path choice estimation). This assumption is reasonable because in most cases there is only one transit route available if $m_i = \text{Transit}$. For driving, biking, or ride-hailing, individuals usually follow the navigation system and thus choose the first option provided by the Google Map API. Note that since individuals with $m_i = \text{Biking}$ do not specify they are using personal bikes or bike-sharing, we assume that they all use personal bikes and ignore the infection risks of touching surfaces.

Another data source we use for infection risk calculation in the transit segment is the smart card data from the Massachusetts Bay Transportation Authority (MBTA). We use the historical smart card data at the same time of the day (one-hour interval) in the last 2 weeks to calculate the mean and variance of the number of passengers in $\mathscr{P}_{s,a}$. Specifically, the smart card data provides the tap-in time and locations. We adopted the destination estimation model proposed by Sánchez-Martínez (2017) to obtain the destination of each trip. Then, a network loading model (Mo et al., 2020) is used to obtain the vehicle load at each time interval. The mean infectious probability $\overline{\mu}_{s,a}$ for passengers in the vehicle is calculated based on regional statistics and their origins. The travel time between stops (i.e., TT_a) is obtained from automated vehicle location (AVL) data.

In terms of the bike and walk segments, there is no open-source pedestrian and cyclist density data available for this study. Therefore, we generate the synthetic pedestrian and cyclists density data by combining population data in Boston (Massachusetts Demographics by Cubit, 2020), Massachusetts Travel Survey (MTS) (Boston Region Metropolitan Planning Organization, 2011), and national household travel survey (NHTS) data (Federal Highway Administration, 2017). The synthetic pedestrian density data include the mean and variance of the number of cyclists and pedestrians at a specific street for each time interval (in this study, every 1 min). The detailed data generation process is shown in A. The basic idea is to generate many bike and walk trajectory samples using the available dataset and aggregate these samples at the street-minute level. Therefore, given a bike/walk segment s of individual i, based on its trajectory, we can sample the number of bike/walk encounters using the generated cyclists and pedestrians density data above. This process is replicated multiple times to get the distribution of $|\mathscr{P}_s|$. The distribution of $\overline{\mu}_s$ is calculated based on neighborhood infection rates (Fisher, 2020). The contact time ($d_{\mathrm{W/B}}$) is assumed to be uniformly distributed with $\mathscr{U}(4,6)$ seconds for a walking trip and $\mathscr{U}(2,4)$ seconds for a biking trip. The breath rates and wind speed parameters for biking and walking are also different (see Table 1 for details).

For the ride-hailing segment, we only need the distribution of $|\mathscr{P}_s|$ (number of passengers), $d_{i,p}$ (shared trip duration). However, there is no public ride-hailing data for Boston. In this study, we use the open-source Transportation Network Company (TNC) data in Chicago (Chicago Data Portal, 2018) to calculate the distribution of $|\mathscr{P}_s|$ and $d_{i,p}$ as an approximation for those of Boston.

Through the Google Map API, we can obtain the value of T_s. However, the distribution of T_s is unknown. Ideally, the distribution of T_s can be obtained from GPS data. Given the data limitations, we assume T_s is normally distributed and the value obtained from the Google Map API is the mean value. The standard deviation is assumed to be $T_s \times 30\%$ based on the empirical study (Li et al., 2013). It is worth noting that, for the transit segment, instead of using Google Map data, we obtain the stop-to-stop travel time using AVL data. Hence, we calculate the segment travel time as $T_s = \sum_{a \in \mathscr{A}_s} TT_a$. The mean and variance of T_s are just the summations of the mean and variance of TT_a, respectively, by assuming independence across segments.

The infection-related parameters are summarized in Table 1. We assume the mask effectiveness parameter to be $\alpha = 50\%$ (Zhou and Koutsopoulos, 2021; Stutt et al., 2020).

4.2. Data description and statistics

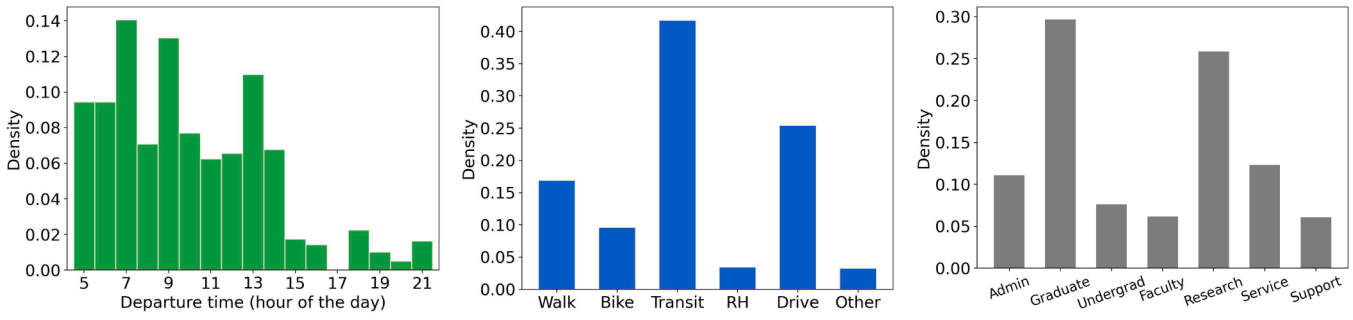
The MIT staff commuting survey consists of 974 individual responses with information on (o_i, d_i, t_i, m_i). Note that only commuting trips toward MIT are considered. The distributions of their departure times, travel modes, and job categories are shown in Fig. 3.

Since only trips toward MIT are considered, most of their departure times are in the morning hours (Fig. 3a). There are also some people going to campus in the evening, which may be students living close by or staff on night shifts. In terms of travel modes (Fig. 3b), more than 40% of the MIT staff and students choose transit as their commuting mode. Driving is the second most popular mode. Of all individuals filling out the survey, around 30% are graduate students. Respondents also include faculty and staff (e.g., admin, service, support, research).

Fig. 4 shows the distribution of travel information collected by the Google Map API. Most of the commuting trips have only one segment (Fig. 4a). The maximum number of segments is five. The travel times of all trips are approximately log-normal distributed with a long tail (Fig. 4b). The travel times for most of the trips are within 1 h.

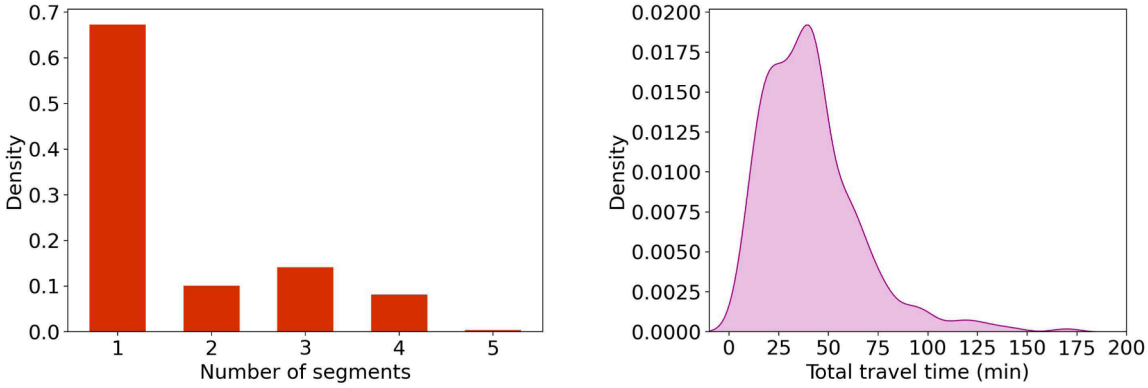
For transit trips, Fig. 5 shows the box plots of travel times between two consecutive stations and the associated passenger load for Bus Route 1 in Boston. Route 1 is a popular bus service along Massachusetts Avenue. The travel time to the first stop is relatively large as vehicles usually depart from garages (Fig. 5a). In terms of passenger load (Fig. 5b), the middle stops in the route have a relatively larger number of passengers onboard. The graphs show that both travel time and passenger loads are uncertain and the uncertainties should be captured in the virus transmission modeling.

Fig. 6 shows the spatial distribution for the inferred number of walk and bike trips at 8:00 AM (details in A). Their spatial distributions are similar. More trips happen at places with higher population densities. The number of walking trips is larger than that of bike trips.



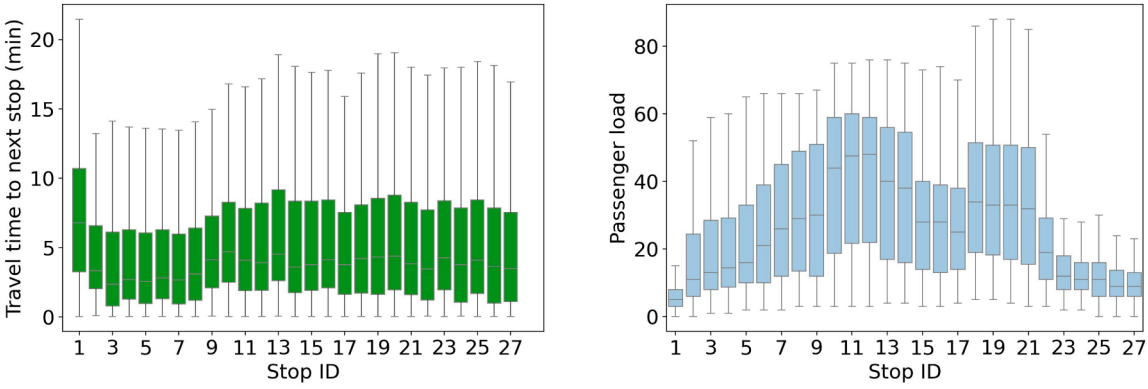
(a) Departure time distribution (b) Travel mode distribution (c) Person type distribution

Fig. 3. Descriptive statistics of MIT staff commuting survey (RH indicates Ride-hailing).



(a) Number of segments distribution (b) Travel mode distribution

Fig. 4. Distribution of the number of segments and travel times.



(a) Travel time to next stop (b) Passenger load

Fig. 5. Box plots of travel time and passenger load for Bus Route 1 (Stop ID represents the stop sequences from north to south).

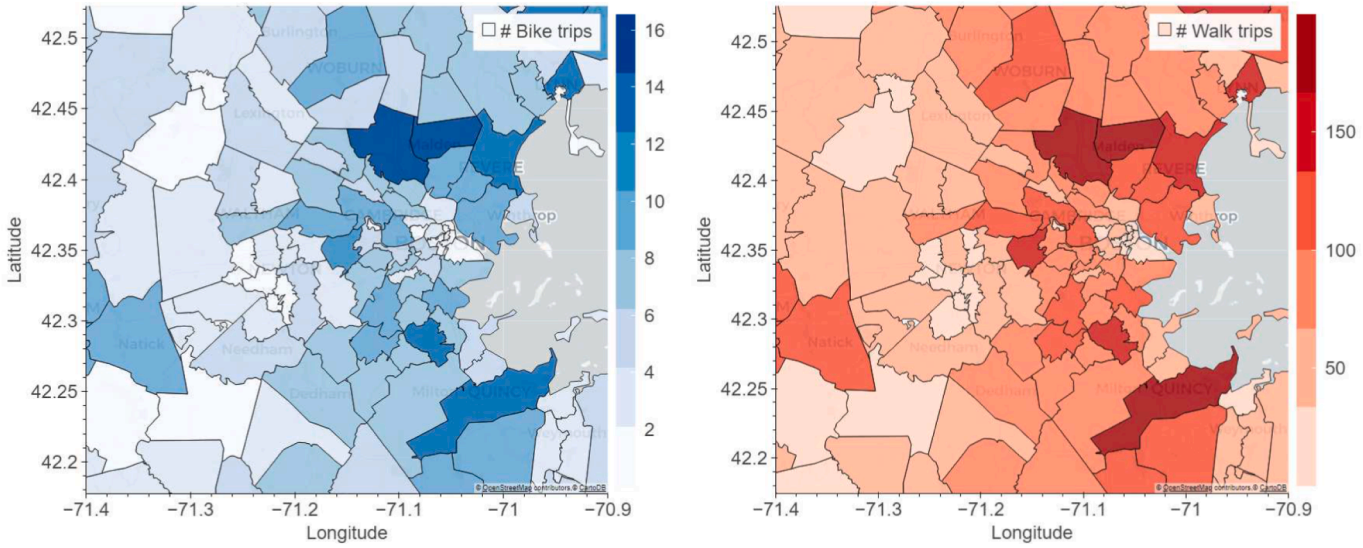
4.3. Results

4.3.1. Infection risk for individuals

After excluding individuals with "other" travel modes. We have 943 remaining individuals. We calculate their infection probabilities and standard deviations using the proposed method. Results are shown in Fig. 7. Most of the individuals have an infection probability close to zero. The maximum infection probability is around 1.3%. This implies that the probability of getting infected during commuting to MIT is quite low. The results are consistent with the previous studies modeling infection risks in transit (Zhou and Koutsopoulos, 2021). In terms of the estimation errors, the standard deviations are approximately 50% of the estimated probability.

Fig. 8 shows the average infection probabilities by travel modes. Ride-hailing, notably, has a similar average infection risk as transit. The potential reason is the low ventilation rates in shared cars compared to transit vehicles. This highlights the importance of epidemic prevention for ride-hailing and other emerging mobility services. The infection risk of transit has a higher variance. This implies that transit may have extremely high risks sometimes when there are many contact passengers and travel time is long.

Since individuals with different travel modes, departure times, and



(a) Number of bike trips (b) Number of walk trips

Fig. 6. Spatial distribution for the number of walk and bike trips at 8:00 AM.

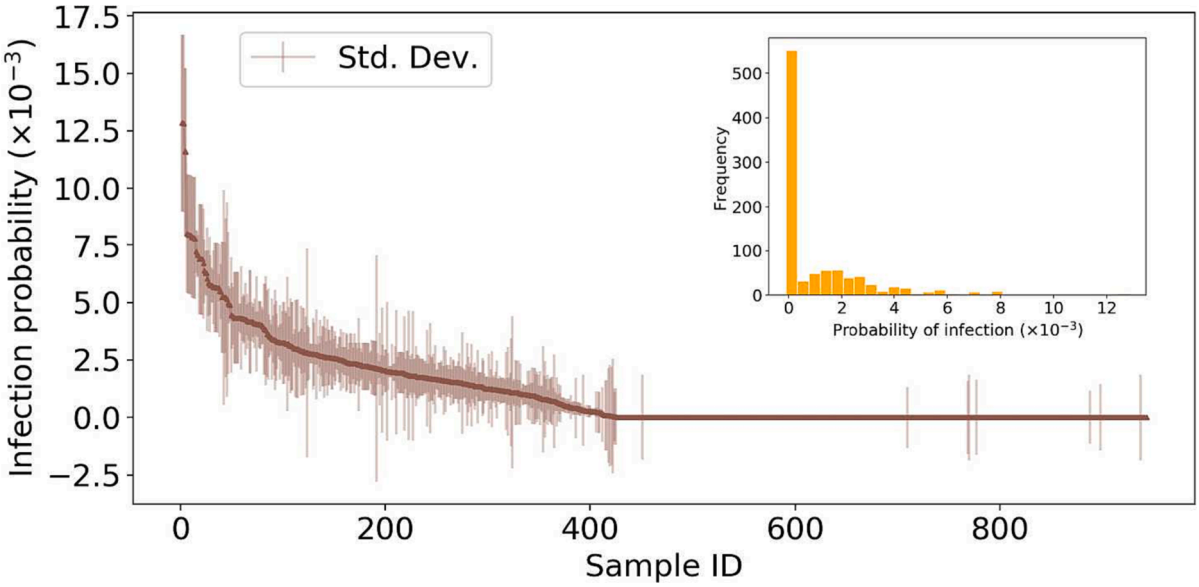


Fig. 7. Inferred individual infection probabilities (sorted by values in descending order).

travel distances may have different infection probabilities. We run a linear regression model to analyze the impact of different factors on infection risks. The dependent variable is the inferred infection probability and the independent variables are as follows:

- If faculty: Whether the individual is a faculty at MIT (Yes $= 1$).
- Distance ($\times$ 100km): Euclidean distance from the individual's home to MIT.
- If transit: Whether the individual's commuting mode is transit (Yes $= 1$).
- If morning peak: Whether the individual's departure time is between 6:00 AM and 9:00 AM (Yes $= 1$).

The results of the linear regression are shown in Table 2. "Distance $\times$ If transit" is added to differentiate the distance impact for transit and non-transit users. We find that transit users are more likely to get infected. And individuals who live further from the school and commute by transit have a higher infection risk. These may be due to the fact that they have a longer travel time and more close contacts, thus are more likely to be exposed to viruses. We also observe people with a departure time in the morning peak have a higher probability of being infected, which may be due to the larger amount of encounters in the morning peak hours. It is worth noting that "Distance" is not significant unless combining them together (i.e., "Distance $\times$ If transit"). This implies that the impact of distance on infection risks mainly applies to transit users.

4.3.2. *Spatiotemporal distribution of infection risk*

In addition to the infection risk calculation for actual survey respondents, the proposed method can also be used to evaluate the spatiotemporal distribution of infection risks by generating synthetic observations.

For the spatial distribution, we generate synthetic observations with home addresses in different neighborhoods. For dummy samples with walking and biking travel modes, we only consider home addresses

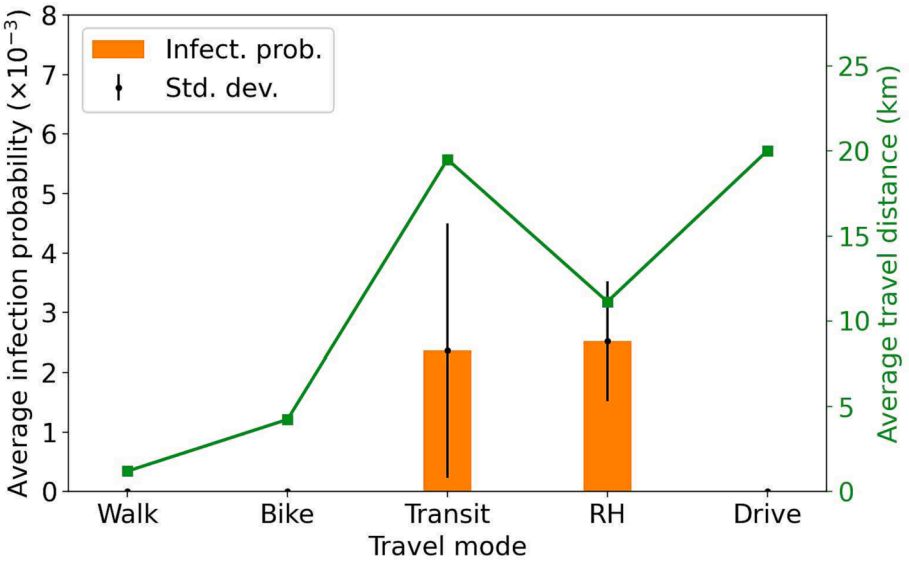


Fig. 8. Inferred individual infection probabilities by travel modes.

Table 2
Factors on infection probability (%).

Variable	Coefficient (Std. Error)	Variable	Coefficients (Std. Error)
Intercept	0.010(0.008)	Distance	0.020(0.044)
If transit	0.159(0.014) **	If morning peak	0.024(0.011) *
If faculty	−0.017 (0.020)	Distance × If transit	0.003 (0.001) **

Number of samples: 943; R^2: 0.386;
**: p-value < 0.05; *: p-value < 0.1.

within 10 km and 20 km of MIT, respectively. All dummy samples' departure times are set at 8:00 AM.

Fig. 9 shows the infection probability of travel mode trips with home addresses in different neighborhoods. In general, people with residences further from MIT have higher infection risks. But the risk is also affected by other factors. For example, transit infection risks are also affected by specific transit routes (i.e., not perfectly proportional to distances). Trips originating from highly populated neighborhoods also have higher infection risks. The scale of infection risks is different, with the highest travel mode of transit, then ride-hailing. The infection risks of biking and walking are much smaller.

In terms of the temporal distribution, we generate synthetic observations with departure times from 0:00 to 24:00. Their home locations are assumed to be uniformly distributed across all neighborhoods. Fig. 10a shows the infection risk as a function of departure time for transit trips. The shaded areas are $\frac{1}{10}$ of the estimation errors. The infection probabilities are higher when people depart during the morning and evening peak hours, which is reasonable because there is usually a higher passenger load (i.e., more close contacts) during rush hours. The infection probabilities for walking trips show a similar pattern (Fig. 10b). The highest infection risks happen in the daytime (from 8:00 to 18:00), most likely due to the larger number of pedestrians.

4.4. Model implementation for return-to-campus policy

Nowadays, many companies (e.g., Amazon (Jassy, 2023)) have required part of their employees to return to the office for specific days of the week. One of the key challenges for these companies is to decide which employees should return to the office. The common way is purely based on commuting distance or time. However, the purely distance-based method may result in high infection risks for some employees during the commuting process. The model proposed in this paper can incorporate infection risks into consideration, so as to allow high infection-risk employees to keep working remotely for health considerations.

We assume that MIT requires at least 50% of their students and employees to return to campus (RTC) and compare a purely commuting distance-based method with a hybrid method considering both commuting distance and infection risks. Specifically, for the distance-based method, we assume MIT selects the top 50% students and employees with the smallest commuting distance. For the hybrid method, we assume the selection is based on a new multi-objective function (denoted as Y_i for individual i):

$$Y_i = D_i + \eta \cdot \mathbb{P}(i \text{ infected} | o_i, d_i, t_i, m_i) \tag{26}$$

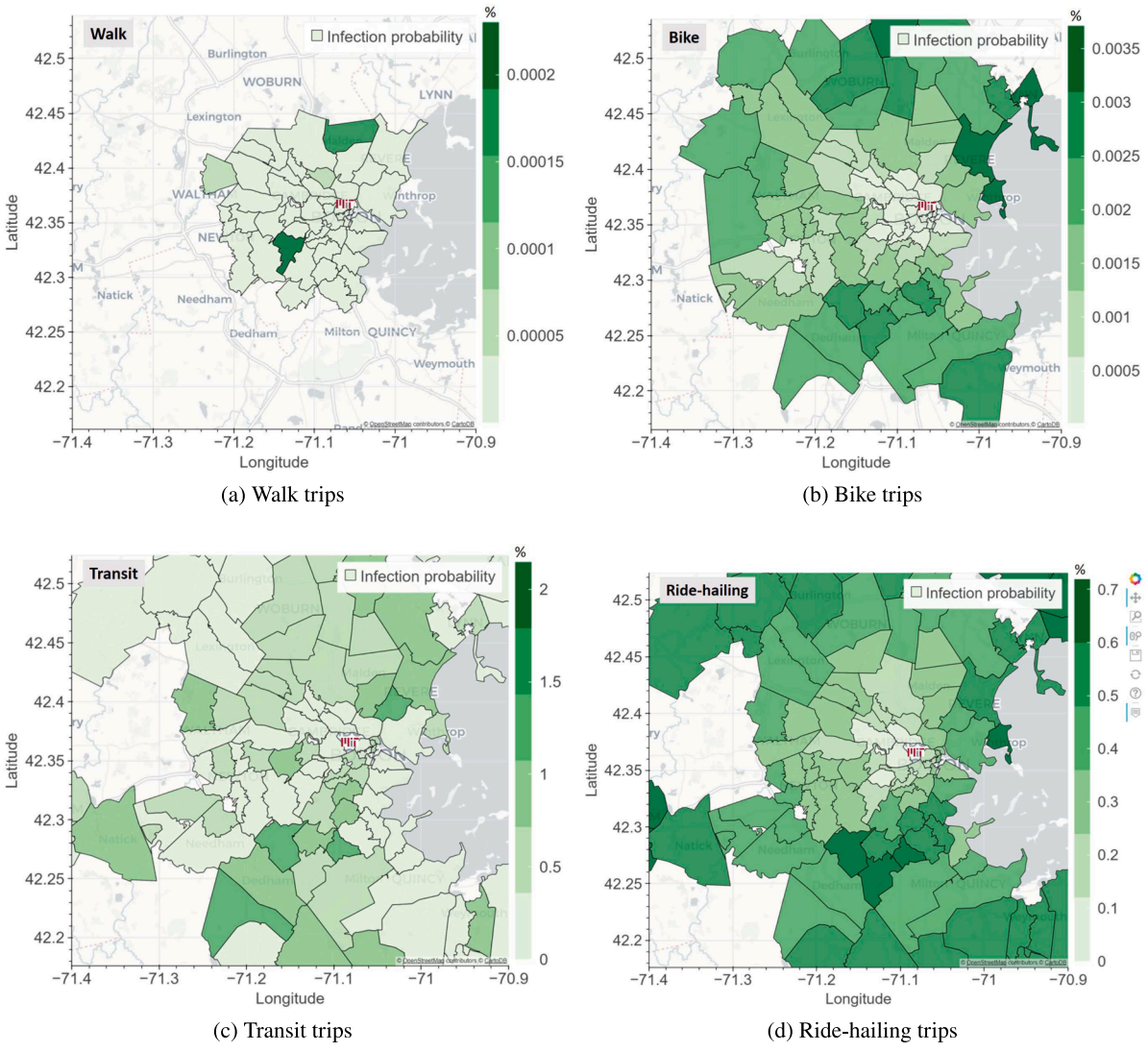
where D_i is the commuting distance of individual i. η is the parameter for the trade-off between commuting distance and infection risks. In this study, we set η to 13254.09, which is $\frac{1}{10}$ of the average distance divided by the average infection probability. Results are shown in Table 3. The model is able to decrease the average infection risks of all RTC employees/students by 38.2% (from 0.06% to 0.037%) while only increasing the average commuting distance by 3.35% (from 4.0 km to 4.13 km).

4.5. Model validation

In this section, we aim to validate that the proposed model can reasonably predict the infection probability during commuting.

4.5.1. Validation with real-world data

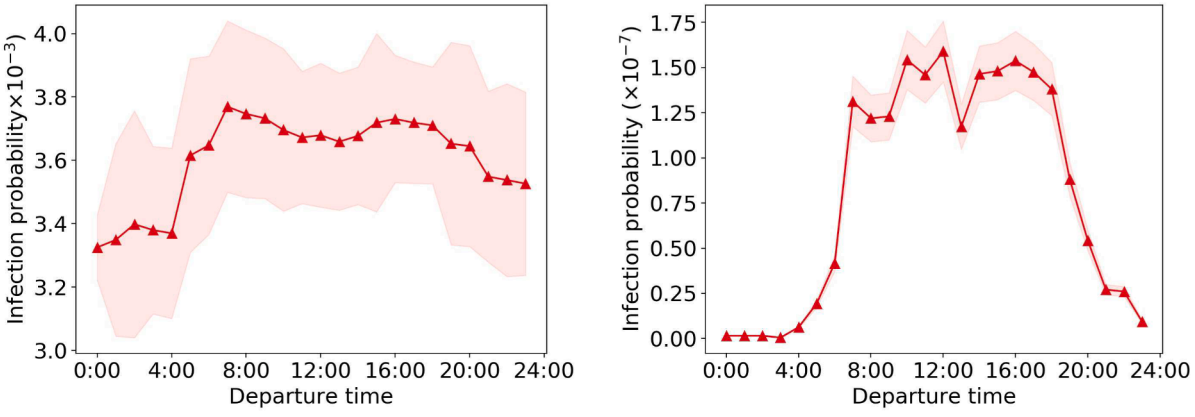
Though there are many reports on infectious cases for different



(a) Walk trips (b) Bike trips

(c) Transit trips (d) Ride-hailing trips

Fig. 9. Spatial distribution of infection probability.



(a) Transit trips (b) Walk trips

Fig. 10. Temporal distribution of infection probability.

communities, these cases are aggregated results from various activities, not just commuting. There are very few studies with real-world infection data that focus on the commuting process. The two case studies we found are 1) Bus infection in Zhejiang, China (Shen et al., 2020) and 2) Bus infection in Hunan, China (Sun and Zhai, 2020).

The parameters results for the two cases are shown in Table 4. These

Table 3
Comparison of different return-to-campus policy.

Method	Num of RTC individuals	Average infection probability (%)	Average commuting distance (km)
Distance-based	472	0.060	4.00
Hybrid	472	0.037	4.13

Table 4
Model validation case studies.

Case	Trip duration	Mask imposed	Ventilation rate	Num passengers	Num of carrier
Shen et al. (2020)	1.7 h	False	$\mathcal{U}(500, 800)$	68	1
Sun and Zhai (2020)	1 h	False	$\mathcal{U}(500, 800)$	12	1

The ventilation rates are obtained from China's standards for metro vehicles GB/T 7928–2003

Table 5
Model validation results.

Case	Actual infected people after trip	Predicted infected people after trip (2 × standard deviation)
Shen et al. (2020)	23	21.03 (10.70 to 31.36)
Sun and Zhai (2020)	3	2.36 (1.12 to 3.60)

cases happen before the locking down of China, where passengers do not wear masks. The comparison results are shown in Table 5. The predicted results are close to the actual infected passengers, validating the effectiveness of the model.

In terms of other transportation modes, to the best of the author's knowledge, there are no real-world infection statistics reported specifically for cycling, walking, or ride-sharing. Hence, the validation for their infection risks is difficult. However, given that the infection mechanism for ride-sharing and transit are similar (i.e., they are both indoor infections modeled by the Wells-Riley model), the validation of buses also provides evidence for the model's effectiveness in ride-sharing. To validate the walking and cycling infection model with real-world data, future studies could consider designing a particle transmission experiment to validate the viral quanta generation and receiving in cycling and walking situations. This is because the Wells-Riley model is built upon virus particle transmission. Another way could be tracking infectious people's daily contacts, selecting individuals who only contact with others during walking and cycling, and calculating the actual infectious probability for these two travel modes.

4.5.2. *Validation with literature*

The Wells-Riley model is adopted in the World Health Organization guidelines for transmission risk control (Chartier and Pessoa-Silva, 2009). The methods are validated and verified in various studies (Foster and Kinzel, 2021; Sun and Zhai, 2020). In this section, we collect different similar studies that calculate the infection probabilities during the commuting process (mostly during transit) and compare the results with ours as an indirect validation. Since the literature data is collected as an aggregated result of all riders, we also show the mean and standard deviation of infection probabilities for all transit users at MIT. Results are shown in Fig. 11. Detailed literature and an explanation of the error bar are shown below:

- Literature [1] (Hu et al., 2021): The values are collected from high-speed train passengers in China (2,334 index patients and 72,093 close contacts). The mean infection rate is 0.32% and the error bar is 95% confidence interval(0.29% - 0.37%).
- Literature [2] (Zhou and Koutsopoulos, 2021): The values are calculated based on the Wells-Riley model using the Boston subway system as a case study. The error bar shows the maximum (0.90%) and minimum (0.49%) average infection risks given different dispatching conditions.
- Literature [3] (Dai and Zhao, 2020): The values are calculated based on the Wells-Riley model. The error bar shows the maximum (0.80%) and minimum (0.40%) values by assuming the same ventilation rate range and travel time as ours.

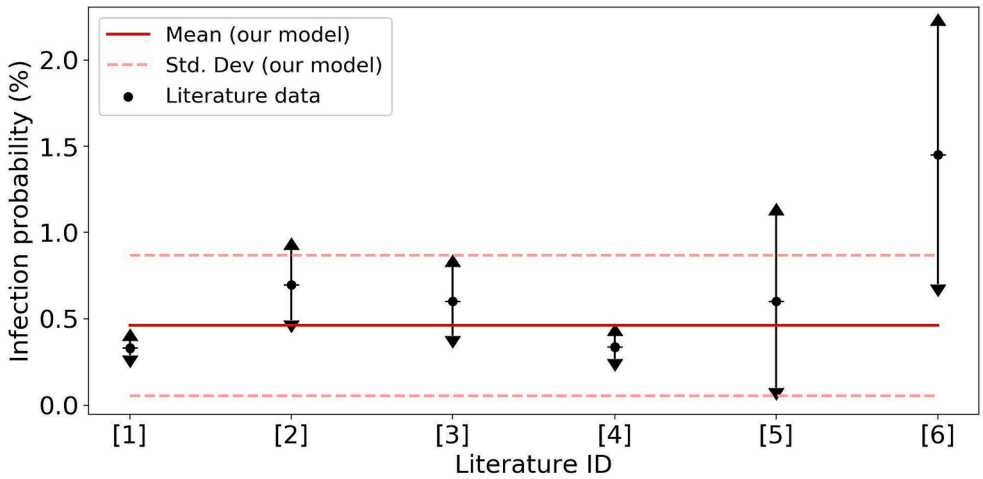


Fig. 11. Comparison between this paper and literature in transit infection.

Table 6
Model validation for ride sharing.

Trip duration	Mask imposed	Ventilation rate	Decayed q_s	Results of the literature (Min, Max)	Our results (2 × std. dev.)
20 min	False	$\mathcal{U}(90,120)$	$\mathcal{U}(23.1,69.2)$	[2.9%, 14.7%]	9.9% [4.6%, 15.2%]

Fig. 12. Example of implementing the model to trip planning.

- Literature [4] (Zhou and Koutsopoulos, 2022): The values are calculated based on the Wells-Riley model using the Boston subway system as a case study. The error bar shows the maximum (0.70%) and minimum (0.24%) values given different dispatching conditions.
- Literature [5] (Shen et al., 2021): The values are calculated based on the Wells-Riley model. The mean infection probability for a subway cabin and bus is 0.6%. The error bar shows the standard deviation (0.5%).
- Literature [6] (Das and Ramachandran, 2021): The values are calculated based on an airborne infection risk estimator developed by Miller et al. (2021). The mean infection probability for a subway cabin and bus is 1.42%. The error bar shows the standard deviation (0.7%).

From Fig. 11, our model results are consistent with those in the literature, which further validate the model.

For other travel modes, Safranek and Scheinker (2022) considers a driver and passenger infection model in a ride-sharing system. Note that Safranek and Scheinker (2022) uses a different infection model which considers the decay of infectivity magnitude by time using a gamma distribution. They also consider two viral variants with different infectious parameters. Therefore, we reduce q_s of our model to the day of symptom onset based on the corresponding gamma distribution. The results are compared with maximum and minimum infectious probabilities in Safranek and Scheinker (2022) on the day of symptom onset. The parameter settings and results comparison are shown in Table 6. We observe a similar infection risk estimation.

Since there are no other studies estimating biking and walking infection probabilities, the comparison of these two modes is not included in this study.

5. Conclusion and discussion

The paper proposes a probabilistic framework to estimate the risk of infection during commuting considering different travel modes, including public transit, ride-share, bike, and walking. The model enables evaluating both the probability of infection and the estimation errors (i.e., uncertainty quantification). Different sources of data (such as smart card data, travel surveys, and population data) are used to estimate commuting individuals' interaction with infectious environments. The model is applied using data related to the MIT community as a case study. We evaluate the commute infection risks for employees and students. Results show that most of the individuals have very low infection probability. The maximum infection probability is around 1.3%. Individuals with larger travel distances, traveling with transit, and traveling at peak hours are more likely to get infected.

The model has several practical applications. 1) The model can be used to support decision-making for companies, schools, or communities during or post the pandemic regarding return to offices. They can collect their employees' commuting information and use the model to evaluate the commuting risk for better planning. For example, employees may use the results to decide which employees should return to the office according to both commuting distance and infection risks. 2) Another implementation of the model is to add individual-level infection risk to their trip planning tool (such as Google Maps). For example, in Fig. 12, the trip planning tool not only shows the recommended routes based on travel time but also the infection risks. Individuals can make better path choice decisions with this additional information. Operators may also provide individual-level path recommendations using both infection risk and travel time information (Mo et al., 2023).

CRediT authorship contribution statement

Baichuan Mo: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Peyman Noursalehi:** Conceptualization, Methodology, Software, Data curation. **Haris N. Koutsopoulos:** Conceptualization, Formal analysis, Writing - review & editing, Supervision. **Jinhua Zhao:** Conceptualization, Formal analysis, Supervision, Project administration, Funding acquisition.

Acknowledgement

The authors would like to thank the MIT Quest for Intelligence for their support and data availability for this research.

Appendix A. Pedestrian and cyclist density calculation

For the infection risk calculation of walking and bike trips, an important input is the number of close-contact encounters during the trip (i.e., the distribution of $|\mathscr{P}_s|$). The distribution can be obtained from the mean and variance of the number of cyclists (denoted as $C_{b,\tau}$) and pedestrians (denoted as $W_{b,\tau}$) at specific street b for each time interval τ. Since there are no GPS or trajectory data available for this study, we generate the distribution of $C_{b,\tau}$ and $W_{b,\tau}$ using the following method.

First, we collect population data (Massachusetts Demographics by Cubit, 2020) for each neighborhood (zip-code level) in the Boston metropolitan area. We use the Massachusetts Travel Survey (MTS) (Boston Region Metropolitan Planning Organization, 2011) to calculate the trip generation rates given the population. As MTS does not include bike and walk mode share, the national household travel survey (NHTS) data (Federal Highway Administration, 2017) is used to get the proportion of bike and walk trips as well as their temporal distributions using samples in Massachusetts.

Combined with the trip generation rate, we simulate the number of bikes and walk trips for each neighborhood at different time intervals. Given limited actual bike and walk trips in NHTS data in Boston, it is hard to obtain the origin–destination (OD) distribution. But the distribution of travel distances and departure times can be obtained. For each bike walk trip, we generate the "hypothetical" trajectory as follows: 1) We first randomly sample an origin within the neighborhood. 2) Then, we sample the travel distance and departure time from the pre-defined distribution, as well as a specific direction (uniformly $0 \sim 2\pi$). The travel distance and direction yield the destination. 3) We generate the trajectory (i.e., a sequence of streets) of the trip. From these trajectories, we obtain the number of pedestrians and cyclists in the street for each street b and time interval τ (i.e., a sample of $C_{b,\tau}$ and $W_{b,\tau}$). This process is repeated multiple times to get $\mathbb{E}[C_{b,\tau}], \mathbb{E}[W_{b,\tau}]$ and $\mathrm{Var}[C_{b,\tau}], \mathrm{Var}[W_{b,\tau}]$.

The generated walking and bike trip distributions are shown in Fig. 6 (the figure is aggregated at the zip code level).

References

Ahangari, S., Chavis, C., Jeihani, M., 2020. Public transit ridership analysis during the covid-19 pandemic. Medrxiv.

Ando, H., Ikegami, K., Nagata, T., Tateishi, S., Eguchi, H., Tsuji, M., Matsuda, S., Fujino, Y., Ogami, A., 2021. Effect of commuting on the risk of covid-19 and covid-19-induced anxiety in japan, december 2020. Arch. Public Health 79, 1–10.

Andrews, J.R., Morrow, C., Wood, R., 2013. Modeling the role of public transportation in sustaining tuberculosis transmission in south africa. Am. J. Epidemiol. 177, 556–561.

Boston Region Metropolitan Planning Organization, 2011. Exploring the 2011 massachusetts travel survey: Mpo travel profiles. URL: https://www.ctps.org/travel-profiles, Last accessed on 2022-11-04.

Buonanno, G., Morawska, L., Stabile, L., 2020. Quantitative assessment of the risk of airborne transmission of sars-cov-2 infection: prospective and retrospective applications. Environ. Int. 145, 106112.

Buonanno, G., Stabile, L., Morawska, L., 2020. Estimation of airborne viral emission: Quanta emission rate of sars-cov-2 for infection risk assessment. Environ. Int. 141, 105794.

Cascetta, E., Nuzzolo, A., Russo, F., Vitetta, A., 1996. A modified logit route choice model overcoming path overlapping problems. specification and some calibration results for interurban networks, in: Transportation and Traffic Theory. Proceedings of The 13th International Symposium On Transportation And Traffic Theory, Lyon, France, 24–26 July 1996.

Chang, H.H., Lee, B., Yang, F.A., Liou, Y.Y., 2021. Does covid-19 affect metro use in taipei? J. Transp. Geogr. 91, 102954.

Chartier, Y., Pessoa-Silva, C., 2009. Natural ventilation for infection control in health-care settings.

Chen, S.C., Liao, C.M., Li, S.S., You, S.H., 2011. A probabilistic transmission model to assess infection risk from mycobacterium tuberculosis in commercial passenger trains. Risk Anal.: Int. J. 31, 930–939.

Chicago Data Portal, 2018. Transportation network providers - trips. URL: https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p, Last accessed on 2022-11-04.

Dai, H., Zhao, B., 2020. Association of the infection probability of covid-19 with ventilation rates in confined spaces. Building simulation, Springer 1321–1327.

Das, D., Ramachandran, G., 2021. Risk analysis of different transport vehicles in india during covid-19 pandemic. Environ. Res. 199, 111268.

Ecke, L., Magdolen, M., Chlond, B., Vortisch, P., 2022. How the covid-19 pandemic changes daily commuting routines–insights from the german mobility panel. Case Stud. Transp. Policy.

Fajgelbaum, P.D., Khandelwal, A., Kim, W., Mantovani, C., Schaal, E., 2021. Optimal lockdown in a commuting network. Am. Econom. Rev.: Insights 3, 503–522.

Federal Highway Administration, 2017. National household travel survey. URL: https://nhts.ornl.gov/, Last accessed on 2022-11-04.

Fennelly, K.P., Nardell, E.A., 1998. The relative efficacy of respirators and room ventilation in preventing occupational tuberculosis. Infection Control Hospital Epidemiol. 19, 754–759.

Fisher, A.R., 2020. Estimated COVID 19 by US Zip Code. URL: https://github.com/adamrfisher/Estimated_COVID19_by_US_Zip_Code, Last accessed on 2022-11-04.

Foster, A., Kinzel, M., 2021. Estimating covid-19 exposure in a classroom setting: A comparison between mathematical and numerical models. Phys. Fluids 33, 021904.

Furuya, H., 2007. Risk of transmission of airborne infection during train commute based on mathematical model. Environ. Health Preventive Med. 12, 78–83.

Gartland, N., Fishwick, D., Coleman, A., Davies, K., Hartwig, A., Johnson, S., Van Tongeren, M., 2022. Transmission and control of sars-cov-2 on ground public transport: a rapid review of the literature up to May 2021. J. Transport Health 101356.

Harvey, A.P., Fuhrmeister, E.R., Cantrell, M.E., Pitol, A.K., Swarthout, J.M., Powers, J.E., Nadimpalli, M.L., Julian, T.R., Pickering, A.J., 2020. Longitudinal monitoring of sars-cov-2 rna on high-touch surfaces in a community setting. Environ. Sci. Technol. Lett. 8, 168–175.

Hu, M., Lin, H., Wang, J., Xu, C., Tatem, A.J., Meng, B., Zhang, X., Liu, Y., Wang, P., Wu, G., et al., 2021. Risk of coronavirus disease 2019 transmission in train passengers: an epidemiological and modeling study. Clin. Infect. Dis. 72, 604–610.

Jassy, A., 2023. Update from Andy Jassy on return to office plans. URL: https://www.aboutamazon.com/news/company-news/andy-jassy-update-on-amazon-return-to-office, Last accessed on 2023-05-13.

Jenelius, E., Cebecauer, M., 2020. Impacts of covid-19 on public transport ridership in sweden: Analysis of ticket validations, sales and passenger counts. Transp. Res. Interdisc. Perspectives 8, 100242.

Ko, G., Thompson, K.M., Nardell, E.A., 2004. Estimation of tuberculosis risk on a commercial airliner. Risk Anal.: Int. J. 24, 379–388.

Kondo, K., 2021. Simulating the impacts of interregional mobility restriction on the spatial spread of covid-19 in japan. Sci. Rep. 11, 1–15.

Ku, D., Yeon, C., Lee, S., Lee, K., Hwang, K., Li, Y.C., Wong, S.C., 2021. Safe traveling in public transport amid covid-19. Sci. Adv. 7, eabg3691.

Li, R., Chai, H., Tang, J., 2013. Empirical study of travel time estimation and reliability. Math. Problems Eng. 2013.

Massachusetts Demographics by Cubit, 2020. Massachusetts zip codes by population. URL: https://www.massachusetts-demographics.com/zip_codes_by_population, Last accessed on 2022-11-04.

Medlock, K.B., Temzelides, T., Hung, S.Y.E., 2021. Covid-19 and the value of safe transport in the united states. Sci. Reports 11, 1–12.

Meredith-Karam, P., Kong, H., Wang, S., Zhao, J., 2021. The relationship between ridehailing and public transit in chicago: A comparison before and after covid-19. J. Transp. Geogr. 97, 103219.

Miller, S.L., Nazaroff, W.W., Jimenez, J.L., Boerstra, A., Buonanno, G., Dancer, S.J., Kurnitski, J., Marr, L.C., Morawska, L., Noakes, C., 2021. Transmission of sars-cov-2 by inhalation of respiratory aerosol in the skagit valley chorale superspreading event. Indoor Air 31, 314–323.

Mitze, T., Kosfeld, R., 2022. The propagation effect of commuting to work in the spatial transmission of covid-19. J. Geogr. Syst. 24, 5–31.

Mo, B., Feng, K., Shen, Y., Tam, C., Li, D., Yin, Y., Zhao, J., 2021. Modeling epidemic spreading through public transit using time-varying encounter network. Transp. Res. Part C: Emerging Technol. 122, 102893.

Mo, B., Koutsopoulos, H.N., Shen, Z.J.M., Zhao, J., 2023a. Individual path recommendation under public transit service disruptions considering behavior uncertainty. arXiv preprint arXiv:2301.00916.

Mo, B., Ma, Z., Koutsopoulos, H.N., Zhao, J., 2020. Capacity-constrained network performance model for urban rail systems. Transp. Res. Rec. 2674, 59–69.

Mo, B., Ma, Z., Koutsopoulos, H.N., Zhao, J., 2023. Ex post path choice estimation for urban rail systems using smart card data: An aggregated time-space hypernetwork approach. Transp. Sci. 57, 313–335.

Mo, B., Von Franque, M.Y., Koutsopoulos, H.N., Attanucci, J.P., Zhao, J., 2022. Impact of unplanned long-term service disruptions on urban public transit systems. IEEE Open J. Intell. Transp. Syst. 3, 551–569.

Ott, W., Klepeis, N., Switzer, P., 2008. Air change rates of motor vehicles and in-vehicle pollutant concentrations from secondhand smoke. J. Exposure Sci. Environ. Epidemiol. 18, 312–325.

Riley, E., Murphy, G., Riley, R., 1978. Airborne spread of measles in a suburban elementary school. Am. J. Epidemiol. 107, 421–432.

Rowe, B.R., Canosa, A., Drouffe, J.M., Mitchell, J., 2021. Simple quantitative assessment of the outdoor versus indoor airborne transmission of viruses and covid-19. Environ. Res. 198, 111189.

Rudnick, S., Milton, D.K., 2003. Risk of indoor airborne infection transmission estimated from carbon dioxide concentration. Indoor Air 13, 237–245.

Safranek, C.W., Scheinker, D., 2022. A computer modeling method to analyze rideshare data for the surveillance of novel strains of sars-cov-2. Ann. Epidemiol. 76, 136–142.

Sánchez-Martínez, G.E., 2017. Inference of public transportation trip destinations by using fare transaction and vehicle location data: Dynamic programming approach. Transp. Res. Rec. 2652, 1–7.

Shen, J., Kong, M., Dong, B., Birnkrant, M.J., Zhang, J., 2021. A systematic approach to estimating the effectiveness of multi-scale iaq strategies for reducing the risk of airborne infection of sars-cov-2. Build. Environ. 200, 107926.

Shen, Y., Li, C., Dong, H., Wang, Z., Martinez, L., Sun, Z., Handel, A., Chen, Z., Chen, E., Ebell, M.H., et al., 2020. Community outbreak investigation of sars-cov-2 transmission among bus riders in eastern china. JAMA Internal Med. 180, 1665–1671.

Shinohara, N., Sakaguchi, J., Kim, H., Kagi, N., Tatsu, K., Mano, H., Iwasaki, Y., Naito, W., 2021. Survey of air exchange rates and evaluation of airborne infection risk of covid-19 on commuter trains. Environ. Int. 157, 106774.

Stephens, B., 2012. Hvac filtration and the wells-riley approach to assessing risks of infectious airborne diseases. National Air Filtration Association (NAFA) Foundation Report.
Stutt, R.O., Retkute, R., Bradley, M., Gilligan, C.A., Colvin, J., 2020. A modelling framework to assess the likely effectiveness of facemasks in combination with 'lock-down'in managing the covid-19 pandemic. Proc. R. Soc. A 476, 20200376.
Sun, C., Zhai, Z., 2020. The efficacy of social distance and ventilation effectiveness in preventing covid-19 transmission. Sustainable Cities Soc. 62, 102390.
Tan, L., Ma, C., 2021. Choice behavior of commuters' rail transit mode during the covid-19 pandemic based on logistic model. J. Traffic Transp. Eng. (English Edition) 8, 186–195.
Wang, H., Noland, R.B., 2021. Bikeshare and subway ridership changes during the covid-19 pandemic in new york city. Transport Policy 106, 262–270.
Wilbur, M., Ayman, A., Ouyang, A., Poon, V., Kabir, R., Vadali, A., Pugliese, P., Freudberg, D., Laszka, A., Dubey, A., 2020. Impact of covid-19 on public transit accessibility and ridership. arXiv preprint arXiv:2008.02413.
Wilson, A.M., Weir, M.H., Bloomfield, S.F., Scott, E.A., Reynolds, K.A., 2021. Modeling covid-19 infection risks for a single hand-to-fomite scenario and potential risk reductions offered by surface disinfection. Am J. Infection Control 49, 846–848.
Zhao, Q., Qi, Y., Wali, M.M., 2022. A method for assessing the covid-19 infection risk of riding public transit. Int. J. Transp. Sci. Technol.
Zhou, J., Koutsopoulos, H.N., 2021. Virus transmission risk in urban rail systems: microscopic simulation-based analysis of spatio-temporal characteristics. Transp. Res. Rec. 2675, 120–132.
Zhou, J., Koutsopoulos, H.N., 2022. Schedule-based analysis of transmission risk in public transportation systems. arXiv preprint arXiv:2202.08505.
Zhu, R., Anselin, L., Batty, M., Kwan, M.P., Chen, M., Luo, W., Cheng, T., Lim, C.K., Santi, P., Cheng, C., et al., 2022. The effects of different travel modes and travel destinations on covid-19 transmission in global cities. Sci. Bull. 67, 588.