Physica A 677 (2025) 130942

Contents lists available at ScienceDirect

Physica A

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

Modeling the alighting and boarding process through train doors using a Markov process with flow-dependent transition probabilities

Mehdi Baali [a,b,*], Christine Buisson [b], Rémi Coulaud [a], Winnie Daamen [c]

[a] Transilien SNCF Voyageurs, 10 rue Camille Moke, Saint-Denis, 93210, France
[b] Université Gustave Eiffel, 25 avenue François Mittérand, Bron, 69500, France
[c] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, CN Delft 2628, The Netherlands

ARTICLE INFO

Keywords:
Railway
Alighting and boarding process
Passenger flow data
Pedestrian flows
Markov process
Fundamental diagram

ABSTRACT

Understanding and modeling the alighting and boarding process in suburban train services is crucial to optimizing train dwellings. The alighting and boarding process is a bi-directional pedestrian flow through a bottleneck, being the door opening. Pedestrian flows, including alighting and boarding processes, are generally modeled by two-dimensional pedestrian models, such as cellular automata or social force models. These two-dimensional models are calibrated from two-dimensional data sources that are often complicated to access for privacy reasons. The availability of disaggregated passenger counting data led us to propose a different modeling approach based on cumulative flows. The model is a Markov process with variable transition probabilities. Transition probabilities are computed from the remaining number of alighting and boarding via a differential equation based on the pedestrian fundamental diagram and density estimations. The parameters of the differential equation were fitted using disaggregated passenger counting data. The model shows better predictive power than a linear benchmark model calibrated on the same data. The physical parameters of the model are consistent with the existing literature. The proposed approach offers an alternative to commonly used two-dimensional models, providing easier calibration. Such a model will enable the forecasting of alighting and boarding time distributions, facilitating better dwell time planning and train and platform design.

1. Introduction

In dense urban areas, railway transportation is increasingly plebiscite to move many people. People go from one station to another, making the train stations places of dense pedestrian flows. Pedestrians have various goals once arriving at a station (leaving the station, going to a platform to wait for a train, etc.) and move through multiple kinds of space (platforms, stairs, tunnels, halls, etc.). The complexity of pedestrian flows needs to be correctly tackled to guarantee safety and comfort for passengers, as well as reliability for train operations. The alighting and boarding process is particularly critical and is unavoidable on any train journey. When large volumes of passengers alight and board, congestion may arise at some station, leading to crowding on the platform, crowding on board, and to dwell time delays [1,2]. These dwell time delays may lead to queuing of trains upstream, named knock-on delays [2,3]. In addition to dwell time delays, congestion during the alighting and boarding process also increases the risk of passenger injuries during the process [4]. Then, alighting and boarding processes must be accurately described to shed light on the

* Corresponding author at: Transilien SNCF Voyageurs, 10 rue Camille Moke, Saint-Denis, 93210, France.
E-mail address: mehdiba5426@gmail.com (M. Baali).

https://doi.org/10.1016/j.physa.2025.130942
Received 16 June 2025; Received in revised form 1 August 2025
Available online 29 August 2025

implementation of congestion control strategies. Therefore, it is necessary to develop precise pedestrian models for the specific case of the alighting and boarding process.

Several approaches have already been proposed to model pedestrian flows. These approaches have been grouped into nine main categories [5,6]. These categories are: continuous models [7,8], cellular automata [9,10], social force models [11,12], velocity-based models [13,14], activity-choice models [15], macroscopic-microscopic hybrid approaches [16], behavioral models [17,18], network models [19,20] and AI approaches [21,22]. From these works, simulators have been built, dedicated to modeling pedestrian flows in various situations. Among them are: Nomad [15], an activity-based model that has been used, for example, to assess the design of a railway station [23], or SimPed [20] used to simulate pedestrians in stairs and tunnels [24], or platforms [25] in railway stations. More recent simulators include, among others, Simba mobi [26], Trajectron++ [27], AgentFormer [28], or MID [29]. Commercial simulators have also been developed. Among them, one may mention VisWalk, MassMotion, or Legion. See, for example, Dubroca-Voisin et al. [30] for a review. VisWalk is based on a social force model, while the other two used activity-choice approaches. The focus of this work is the alighting and boarding process.

From a pedestrian point of view, the alighting and boarding process consists of passing through a train door (consisting of a bottleneck with a limited width) in a bi-directional flow. For this more specific case, some of the above-mentioned approaches have been applied to simulate alighting and boarding flows. In general, social force models [31–36] or cellular automata [37–39] are chosen to model the alighting and boarding process. For example, Li et al. [36] derived a social force model with particular attention to the influence of people boarding before all alighting people had alighted (known as non-compliance behavior). Zhang et al. [37] used a cellular automata model to consider individual differences to model the process. The alighting and boarding process is limited in time and space, which makes the dynamics insensitive to strategic and tactical pedestrian choices. Activity-choice models, velocity-based models, or behavioral models are then less suitable for alighting and boarding simulations.

In addition, like any human behavior phenomenon, the alighting and boarding process is subject to important variability, which needs to be modeled to manage dwell time and prevent dwell time delays. Buchmueller et al. [40], for example, measured the different sub-processes of the dwell time from Swiss field data and showed that most of the variability in dwell times came from the alighting and boarding processes (even with a given number of movements). Similar results have been obtained by Seriani et al. [41] in a laboratory experiment. Both have highlighted non-trivial variability patterns: Buchmueller et al. [40] have seen that alighting and boarding time variability increases more slowly than the number of passenger movements. Seriani et al. [41] have obtained decreasing per-passenger variability with time. However, none of these references have commented much on these variability patterns. Every individual is different from one another (inter-individual variability), and people's behaviors may change according to the context (intra-individual variability). This variability is inherent in the pedestrian flows involved. Therefore, microscopic approaches such as social force models or cellular automata have generally been researchers' choices to simulate the variability of the flow. Notice that network-based models can also consider stochasticity in the process.

However, pedestrian approaches to model the alighting and boarding process have rarely been calibrated with precise data. Some models are not calibrated from data at all, using parameters from the literature to study specific influences of certain features [35,36,38]. Some models are calibrated on aggregated data, generally alighting and boarding times [33,34,37]. When calibrated with precise data, the most common choice is video data [31,32,39], as discussed in the review by Haghani and Sarvi [42].

At the same time, it is sometimes difficult to access and use video cameras for privacy reasons. Thus, we propose a novel approach to model the alighting and boarding process based on passenger counting data that are increasingly available in large quantities in some train lines. Several works have already used this kind of data for modeling purposes, but, to our knowledge, only to model minimum dwell times, or alighting and boarding times [43,44]. Contrary to existing approaches, which are based on two-dimensional representations of pedestrians around the train door, we propose to model the alighting and boarding process from cumulative flows. Cumulative flows have already been discussed for pedestrians in Daamen et al. [45], for example. Our goal is to use passenger counting data at the most disaggregated scale possible (here, at the door counting event scale, number of people that have passed every several seconds), to calibrate a cumulative flow model of the alighting and boarding process. This level of resolution for real-world data has rarely been used so far, with an exception in the study of bus alighting and boarding times using automatic fare counting data [46]. Disaggregated counting data provides information on the time each person crosses the door, but no information on the spatial distribution of the passengers on both sides of the door. This information is, yet, crucial for the existing approaches. Describing the alighting and boarding process from this piece of data then requires an innovative approach.

To our knowledge, only one kind of model has been proposed so far in the literature to model pedestrian cumulative flows: queuing models. This is, for example, the case in the work of Løvås [19] to get from one room to another in a network-based model. However, Løvås [19] assumes constant queue rates (independent of any flow or density condition). This kind of model will serve as a benchmark in our study because it can be fitted to our real-world data. In contrast, our cumulative flow model will also consider density conditions.

The flow-density relationship for pedestrians is described by so-called fundamental diagrams. The oldest example for pedestrians is probably [47], but the concept had been used for a long time for cars [48]. Many relationships have been proposed since [49, 50]. Vanumu et al. [51] proposed a review of existing fundamental diagrams in various contexts (uni-directional and bi-directional flows, bottlenecks, stairs, etc.) and discussed the resulting differences. For alighting and boarding processes, the study of Daly et al. [52] is an example of a fundamental diagram plot from experimental data.

In this work, we will use disaggregated passenger counting data to fit varying alighting and boarding rates depending on estimated density conditions. The fitted values will serve as an average rate for a stochastic queuing model to represent the cumulative flow of alighting and boarding. This work aims to model pedestrian flows at train doors with an original stochastic

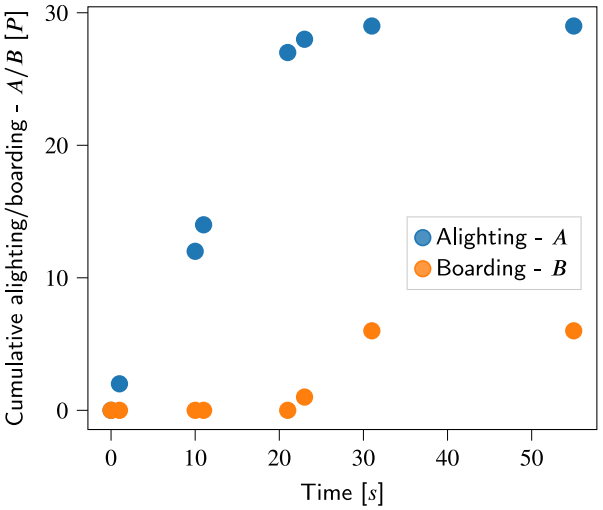

Fig. 1. Example of cumulative data provided by the infrared sensors for one stop and one door. Time 0 corresponds to the end of the opening of the door.

approach: a queuing model with varying alighting and boarding rates depending on the density conditions. The model is fitted to disaggregated counting data.

In summary, our contributions are:

- Proposing a cumulative flow deterministic model calibrated on disaggregated passenger counting data to model the alighting and boarding process,
- Proposing a stochastic cumulative flow model that is a generalization of the deterministic model,
- Estimating the alighting and boarding times distribution using the stochastic model.

First, we precisely present the data source, the data selection, and the delimitation of the perimeter. Next, the method is presented, going from the general idea to the precise formulas used here. The following section is dedicated to the results showing the good ability of the model to describe the alighting and boarding process and the alighting and boarding time compared to a linear benchmark from a 5-fold cross-validation computation. The physical meanings of the parameters are also interpreted. This article ends with a conclusion.

2. Data selection and perimeter

In this work, we propose a cumulative flow model for the alighting and boarding process. One of the major innovations of our approach is the use of disaggregated passenger counting data to calibrate and validate the model. This data is first introduced before describing the model in further detail in Section 3.

Disaggregated passenger counting data is provided by infrared sensors above each train door that count the cumulative numbers of alightings and boardings. This technology is available for a new rolling stock called Regio2N. Every few seconds, information on the cumulative number of alightings $A(t)$ along with the cumulative number of boardings $B(t)$ is registered (see Fig. 1). We name this information a counting event. The time separating two counting events is truly random with an average of 3s and a standard deviation of 3s. The system also registers for each door whether the door is open or closed. A study from the data provider assessed the quality of the sensors using manual counting. They found a root mean square deviation of 4.4 people boarding and 3.6 people alighting. Such deviation is quite high, but mainly corresponds to variance, having a bias of +0.13P in alighting, and +0.24P in boarding. The amount of data should then mitigate the effect of data quality in the fitting. We will confirm this statement in Section 4. The dataset is over six months (September 2022 to February 2023), where we have data for almost 30,000 stops (see Table 2).

Each unit of the Regio2N (see Fig. 2) is an articulated train composed of eight cars. Not all cars are fully similar:

- cars 3, 5, and 7 are identical, contain two doors, and are single-deck;
- cars 2, 4, and 6 are identical, do not contain doors, but are double-deck;
- cars 1 and 8 contain one door and are not identical.

Furthermore, except for cars 1 and 8, the cars are symmetric. In addition, the eight doors of the train are 1.60m wide, and the doors of cars 1 and 8 (that is, door 1 and door 8) are equipped with sliding platforms. Existing works on alighting and boarding times showed a dependency on the train design. Harris [53] considered the width of the door and the number of seats in the carriage. Su

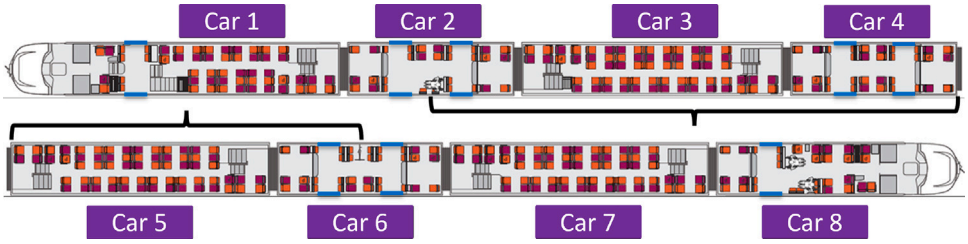

Fig. 2. Design of the Regio2N rolling stock. Parts marked with braces are selected.

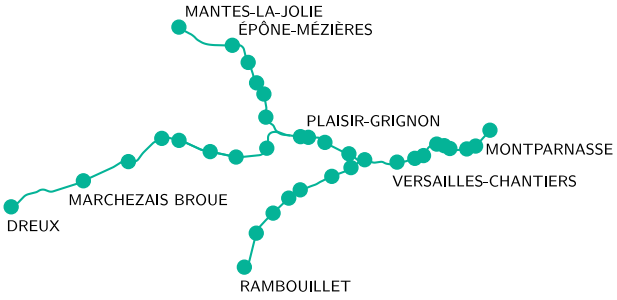

Fig. 3. Map of line N. Commuter train line in the Paris (France) suburban area.

et al. [34] showed an influence on the proportion of seats inside the train based on simulation. The influence of door width is regularly highlighted in the literature on bottlenecks [54], and an example for buses is also provided by Fernández et al. [55]. More generally, the rolling stock design is considered a variable influencing the alighting and boarding process [56,57]. Qualitatively, the interior design should not impact the flow under low to moderate passenger volumes, as the vestibules inside the train are the same for all doors. However, as a precaution, we only considered doors 3 to 6, where the interior designs related to them are symmetrically similar. Moreover, Daamen et al. [58] showed an influence of vertical and horizontal gaps between the train and the platform. Doors 3 to 6 are identical in this respect and without sliding platforms.

Line N of the Paris suburban network is operated only by Regio2N. Fig. 3 shows the map of this line, from which the data for this study come. The platforms in this line are the same height as the train floor, except for platforms at *Marchezais Broué* and *Epône-Mézière*, where it is lower than the other stations. Both stations were excluded due to the influence of vertical gaps on the alighting and boarding process [58]. The influence of vertical gaps or steps was also proven for buses [55,59]. If many stations have lower heights, the model can be calibrated separately on these stations. In our case, only two stations are concerned, so we chose to remove them.

Line N is a commuter line: it is unlikely to see people with lots of luggage, whose influence (shown by Daamen et al. [58] or Kuipers et al. [60]) can, thus, be neglected. The volume of passengers in this line is intermediate, with an average load of 169 passengers, much less than the capacity (1054 passengers). Furthermore, the seating capacity (524 passengers) is exceeded 0.2% of the time. Then, we can neglect the impact of on-board passengers (neither alighting nor boarding), though proven qualitatively important for alighting and boarding modeling considering congestion levels [41,61], and quantitatively considering their number [43,53].

In this perimeter, we consider three situations: only alighting, only boarding, and bi-directional flows. The total number of alighting and boarding at each door is a widespread variable in modeling alighting and boarding times [43,44,53] and alighting and boarding rates [60,62]. The relationship with the total number of alighting and boarding often includes non-linearities [43,53,60,62], or dependencies of the parameter with the density around the door [63]. The relationship also often considers the boarding over alighting ratio [41,64]. For example, in non-crowded situations [64] showed significantly higher alighting and boarding rates with ratios of 0.25 or 4 compared to a ratio of 1 when passengers complied with the alighting priority. In contrast, the opposite result was obtained when passengers did not comply with the alighting priority. For a first approximation, we will consider any passenger movement (alighting or boarding) without distinction in the bi-directional case and neglect the boarding/alighting ratio.

In each situation, the number of movements for a given door is rarely higher than 20 movements; see Fig. 4. With such volumes, no congestion is likely to happen in the train or on the platform. In particular, the alighting area on the platform does not influence the alighting and boarding process, as shown by Yang et al. [33] from simulations.

Furthermore, we consider only the alighting and boarding process with more than six alighting people or boarding people in the case of uni-directional flows, more than six movements (alighting plus boarding) with at least one alighting and one boarding in the case of bi-directional flows. On the one hand, having too few passengers involved in the process may give too much weight to their inter-individual differences and variability in their behavior, let alone sensors' errors. On the other hand, we need to have enough events in the dataset for the study. The number six is a trade-off for a decent consideration of both restrictions. The vast

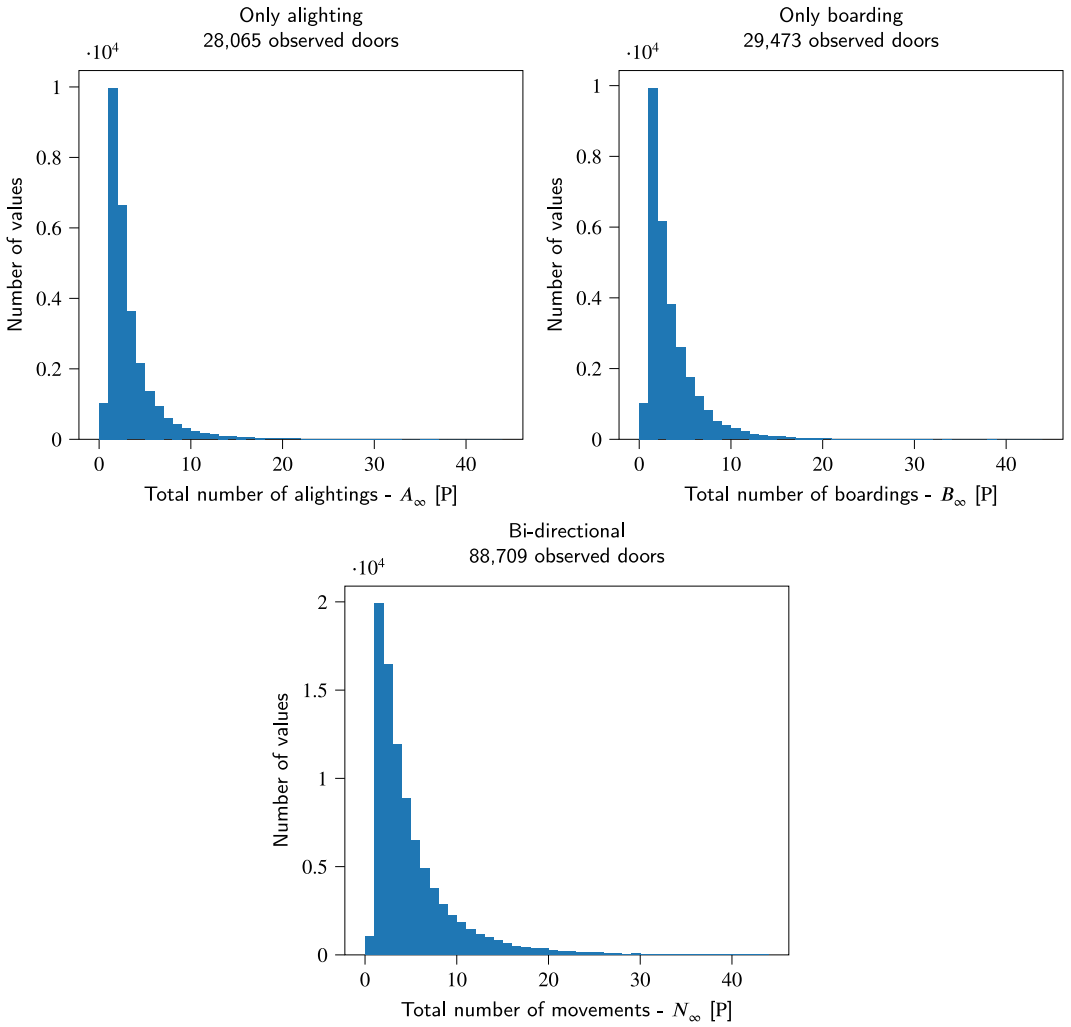

Fig. 4. Histograms of total numbers of alightings (top left), boardings (top right), and movements (bottom) for uni-directional flows (top) and bi-directional flows (bottom) for any open door. Most alighting and boarding processes are composed of only a few movements.

majority of observations were with fewer than six movements. Keeping them for calibration may bias the model for larger numbers of movements. This threshold of six can be found using an elbow criterion (see Fig. 4). At the same time, from a practical application point of view, the situations that need to be properly modeled are situations at higher volumes (where congestion or dwell time delays may occur).

Passengers have several kinds of inter-individual differences that may impact the alighting and boarding process modeling. For example, passenger ages [59], genders [37], or body shapes [35] were shown to influence the process. Other aspects related to variability in passengers' behaviors were studied in the literature, such as passengers' ability to compete to get in or out of the train [37] or their individual desired speed [35]. All the variability due to passengers' differences or behaviors will be considered as random in this work. More details on how we model randomness will be provided in Section 3.

Among the remaining data, other sub-samples will be considered on which the model will also be calibrated to analyze the effect of the following characteristics:

- lower non-compliance and higher non-compliance to alighting priority, shown to have a weak but non-negligible effect [65,66] (we consider higher non-compliance when at least 3 boarding passengers board before 90% of alighting people have alighted, lower non-compliance for 2 or fewer non-compliant boardings);
- peak hours (6 am to 10 am, 4 pm to 8 pm) and off-peak hours, where it is likely to see different patterns occurring, but not always observed [57];

Table 1
Features determining the alighting and boarding process considered in the literature and their treatment in our approach. Five sub-samples are considered to analyze the impact of some characteristic on the calibrated parameters. These sub-samples are respectively peak hours, off-peak hours, lower non-compliance, higher non-compliance, and the specific station of *Versailles-Chantiers*.

Feature	Reference	This approach				
		Variable	Constant	Neglected	Random	Sub-sampled
Demand						
Alighting	Harris [53] and Coulaud et al. [44]	X				
Boarding	Harris [53] and Coulaud et al. [44]	X				
On-board	Harris [53] and Cornet et al. [43]			X		
Flow features						
On-board congestion	Seriani and Fernandez [61] and Seriani et al. [41]			X		
Density around the door	Zhang et al. [63]	X				
Non-linearities	Cornet et al. [43] and Kuipers et al. [60]	X				
Boarding/alighting ratio	Seriani et al. [41] and Fu et al. [64]				X	
Rolling stock features						
Number of seats	Harris [53] and Su et al. [34]		X			
Effective door width	Harris [53] and Fernández et al. [55]		X			
Rolling stock type	Thoreau et al. [56] and Christoforou et al. [57]		X			
Platform features						
Horizontal gap	Daamen et al. [58]		X			
Vertical gap	Daamen et al. [58] and Fernández et al. [55]		X			
Alighting area width	Yang et al. [33]			X		
Station	Qu et al. [67] and Christoforou et al. [57]					X
Passenger-related features						
Peak/off-peak hour	Christoforou et al. [57]					X
Non-compliance	Wahaballa et al. [65] and Seriani et al. [66]					X
Luggage	Daamen et al. [58] and Kuipers et al. [60]			X		
Ages	Tirachini [59]				X	
Genders	Zhang et al. [37]				X	
Body shapes	Yang et al. [35]				X	
Competition	Zhang et al. [37]				X	
Desired speed	Yang et al. [35]				X	

- the specific station of *Versailles-Chantiers*, main station in *line N* except for the Paris terminus, where differences may occur due to station differences or higher demand on average [57,67].

Several determinants of the alighting and boarding process discussed in the literature are constant in the entire perimeter or specific sub-samples of the data. Table 1 summarizes them. Other factors will be considered as variables in the model or as random.

Table 2 summarizes the different steps of the data selection and the studied sub-samples of data, along with the remaining data quantities, average number of movements, and average flows. Only a little variation is observed in the average numbers of movements and flows between the studied sub-samples.

3. Method

As already mentioned, this work aims to propose a model describing the alighting and boarding process from a cumulative flow approach. Formally, we aim to get the dynamics of the cumulative alighting $A(t)$ versus time, the cumulative boarding $B(t)$ versus time, and the cumulative passenger movements $N(t)$ versus time at the door level. First, we derive a differential equation to get a deterministic model, which is the average dynamics of the stochastic model. Then, we include randomness, considering the alighting and boarding process as a Markov process (with varying rates depending on the estimated density condition). Both the deterministic part and the stochastic part will be assessed separately.

3.1. Deterministic modeling of the alighting and boarding process

First, we derive deterministic differential equations to obtain the average dynamics of the alighting and boarding process. Our approach differs from macroscopic two-dimensional approaches in several ways, which we will detail in this section. Table 3 summarizes them.

3.1.1. General differential equation system

For the sake of simplicity, we will note X any of A, B, or N. For better readability, we will not write all the indices concerning the stop and the door (train number, station, date, door number), and only write $X(t)$. $t = 0$ is set at the end of the door opening.

Table 2
Summary of the steps in the data selection and the studied sub-samples, along with the remaining data quantities and the average number of movements and flow. Only a little variation is observed in the average numbers of movements and flows between the studied sub-samples. Notice that a stop is not considered if all remaining doors have 0 alightings and 0 boardings.

Selection step or sub-sample	Number of stops	Number of counting events	Average number of movements [P]	Average flow [$P \cdot s^{-1}$]
All counting data on *line N*	29,768	764,263	4.87	0.41
Only doors 3 to 6	27,839	379,725	4.59	0.40
All stations but *Marchezais Broué* and *Epône-Mézière*	27,357	375,465	4.61	0.40
More than six people alighting (unidirectional case)	2347	14,892	9.10	0.52
More than six people boarding (unidirectional case)	2634	19,752	8.87	0.40
More than six people of both (bidirectional case)	8347	98,841	10.20	0.58
Peak hours	4347	53,744	11.71	0.60
Off-peak hours	4000	45,097	10.60	0.57
Lower non-compliance	7955	85,127	10.88	0.59
Higher non-compliance	1606	13,714	13.70	0.56
Only *Versailles-Chantiers*	1517	27,928	13.77	0.63

Table 3
Differences between our approach and a two-dimensional macroscopic approach.

	This approach	Macroscopic 2D approach
Use case	Crossing (here train doors)	General
Data	High-resolution counting	Video
Variables	Cumulative flows	Flows & Densities
Access to density	Implicit	Explicit
Equation associated with fundamental diagram	Density function	Conservation law

The boundary conditions are then $X(t=0)=0$ and $X(t=+\infty)=X_\infty$, where X_∞ is the total number of alighting people, or boarding people, or movements at this door and this stop.

A traditional way to analytically describe pedestrian flows is to use a fundamental diagram and a conservation law. The main assumption behind our approach, a cumulative flow approach, is that the pedestrian speed (v) only depends on the pedestrian density (ρ), $v = v(\rho)$. ρ will then be estimated from another equation. From this, in a time-dependent situation, the pedestrian flow (ϕ) is given by $\phi(t) = \rho(t)v(\rho(t))$. A conservation law is not relevant in our case since it is not a two-dimensional problem, but only a cumulative flow model. Instead, we derive a differential equation for X to adapt this relation to the cumulative flow modeling. The density is a hidden variable in the final equation. It can be interpreted as the average density in the vicinity of the door (more specifically in the vestibule for the only alighting case, on the near platform for the only boarding case, and both sides for bi-directional flows). This interpretation is similar to the notion of density provided by Zhang et al. [63].

We have, from the definition of cumulative flows: $D\phi(t) = \frac{dX}{dt}$ (where D is the door width). This means that $\frac{dX}{dt} = D\rho(t)v(\rho(t))$. Densities should vary with time. In particular, as people alight/board, the pressure inside/outside the train should decrease, leading to a decrease in density. Thus, we postulate that $\rho(t)$ only depends on $X(t)$: $\rho(t) = f(X_\infty - X(t))$. f is the density function that serves to estimate ρ from X. As a reminder, X_∞ is the total number of movements. Then, the density at time t is assumed to depend only on the number of movements remaining at time t: $X_\infty - X(t)$.

This leads to an ordinary differential equation:

$$\frac{dX}{dt} = Df(X_\infty - X(t))v(f(X_\infty - X(t))). \tag{1}$$

Table 4 summarizes the important variables in this work.

One may object that the alighting and boarding process is a bi-directional flow. Yet, most of the time, fundamental diagrams for bi-directional flows consider the total flow (sum of the flow from both directions) versus the total density (density from both directions) as shown in the review of Vanumu et al. [51]. In the alighting and boarding process, a common rule is to give priority to alighting. Some boarding passengers may not comply with this rule, and such behavior cannot be modeled with flows and densities. However, Baali et al. [68] shows that compliance with this rule has little impact on cumulative movement dynamics ($N = A + B$) with low to intermediate total movements and a balanced boarding/alighting ratio overall. This counterintuitive result is due to similar flows between a density below the optimal density (density resulting in the maximal flow) with lower compliance and a density higher than the optimal density with higher non-compliance. The approach proposed in this work cannot predict whether and when non-compliance behavior may occur and, consequently, the dynamics of the boarding/alighting ratio versus time. Then,

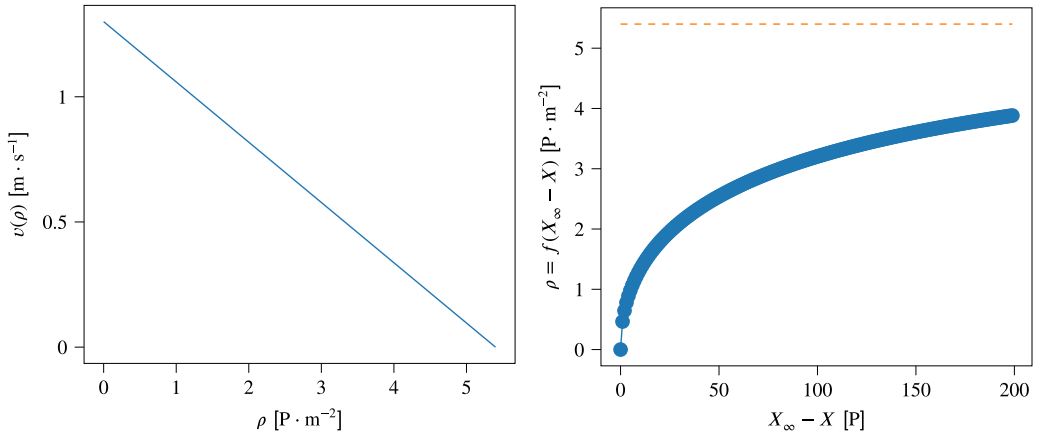

Fig. 5. Illustration of the chosen fundamental diagram (left), illustration of the chosen density function (right), for $X_\infty = 200$P.

we will neglect the effect of both non-compliance and boarding/alighting ratio and keep the same form of ordinary differential equation for N.

In the next subsection, the speed function (v) will be chosen among the fundamental relations proposed in the literature, and the density function (f) will be conjectured from the previous assumptions. v and f are assumed to have the same form and the same parameters throughout the dataset. The differential equation obtained will be solved to obtain A, B, or N versus t.

3.1.2. Choices for evolution equations

As presented before, two relations have to be set to complete the differential equation: the fundamental diagram and the density function. This work aims to assess the method presented above, using one example for both of these relations. The choice of these relations is qualitatively discussed, and future work should explore other forms of relations.

Fundamental diagram Concerning the speed function of the fundamental diagram, we select the Greenshields linear relation, as often proposed in the pedestrian literature [69,70]. Its expression is the following:

$$v(\rho) = v_0 \left(1 - \frac{\rho}{\rho_{max}}\right), \tag{2}$$

where v_0 is the free-flow speed, ρ_{max} is the maximal density. An illustration of this relation is presented in the left part of Fig. 5. Other formulations have been proposed in the pedestrian literature, such as Underwood formulations [71,72], piecewise formulations [73], or non-linear formulations with additional parameters [49]. The Greenshields formulation is a linear relationship with two physically interpretable parameters. This simple formulation was shown to be inappropriate for high densities [74], but from the volumes observed in the perimeter studied, it is unlikely that high densities occur. Then, a Greenshields formulation would not lead to significant inaccuracies, and no data with sufficient volume would be able to fit high densities. Therefore, we focus only on the Greenshields formulation.

Notice that even if no congested regime is likely to be observed in practice, a fundamental diagram is still justified. Before the critical density, the fundamental diagram describes a non-linear increase in flow with the density that is important to be reproduced by the model.

Density function Concerning the density function f, we set four assumptions.

1. The more people there are left to alight/board, the higher the density. Then f must be increasing.
2. A maximum density ρ_{max} cannot be exceeded, and the space in the train/the platform is limited. Then, f should tend to this maximal density, $\lim_{X_\infty - X \to +\infty} f(X_\infty - X) = \rho_{max}$.
3. For no people, the density is zero: $f(0) = 0$.
4. When there are few people, people tend to accumulate around the door, increasing the density faster. Then, we assume f locally strictly concave close to 0: $\frac{d^2 f}{d(X_\infty - X)^2} > 0$. This assumption is inspired by the observations made by Wu and Ma [75] concerning the positioning of boarding people on the platform.

Given these assumptions, we propose the following form for the density function:

$$f(X_\infty - X(t)) = \rho_{max} \left(1 - \exp(-\gamma \cdot \sqrt{X_\infty - X(t)})\right), \tag{3}$$

where ρ_{max} is the maximal density, and γ is a density calibration parameter. An illustration is presented on the right part of Fig. 5.

Table 4
Summary of variables and their units. Notice that the transition rate is a variable only used for the stochastic model in Section 3.2.

Name/Expression	Symbol	Unit	Type
Cumulative alighting	A	[P]	Variable
Cumulative boarding	B	[P]	Variable
Cumulative sum of alighting and boarding	N	[P]	Variable
Total alighting	A_∞	[P]	Variable
Total boarding	B_∞	[P]	Variable
Total sum alighting and boarding	N_∞	[P]	Variable
Replace A, B, or N	X	[P]	Variable
Flow	ϕ	$[\mathrm{P}\cdot\mathrm{m}^{-1}\cdot\mathrm{s}^{-1}]$	Variable
Density	ρ	$[\mathrm{P}\cdot\mathrm{m}^{-2}]$	Hidden variable
Speed	v	$[\mathrm{m}\cdot\mathrm{s}^{-1}]$	Hidden variable
Free flow speed	v_0	$[\mathrm{m}\cdot\mathrm{s}^{-1}]$	Hidden parameter
Maximal density	ρ_{max}	$[\mathrm{P}\cdot\mathrm{m}^{-2}]$	Hidden parameter
Door width (1.6 m)	D	[m]	Fixed parameter
Density calibration parameter	γ	$[\mathrm{P}^{\frac{1}{2}}]$	Fitted parameter
Flow size ($v_0 D\rho_{max}$)	ψ	$[\mathrm{P}\cdot\mathrm{s}^{-1}]$	Fitted parameter
Transition rate to state i	μ_i	$[\mathrm{P}\cdot\mathrm{s}^{-1}]$	Variable

Obtained differential equation The obtained differential equation is therefore:

$$\frac{dX}{dt} = Dv_{0,X}\rho_{max,X}\left(1-\exp(-\gamma_X\cdot\sqrt{X_\infty - X(t)})\right)\left(\exp(-\gamma_X\cdot\sqrt{X_\infty - X(t)})\right), \tag{4}$$

see Table 4 for a reminder of variables and parameter definitions.

This equation has only two degrees of freedom ($\psi_X = Dv_{0,X}\rho_{max,X}$ and γ_X), it is then more appropriate to write:

$$\frac{dX}{dt} = \psi_X\left(1-\exp(-\gamma_X\cdot\sqrt{X_\infty - X(t)})\right)\left(\exp(-\gamma_X\cdot\sqrt{X_\infty - X(t)})\right), \tag{5}$$

ψ_X can be interpreted as a flow size.

3.1.3. Parameters fitting

From the equation presented above, we will get the best parameters from a fit. A traditional quadratic cost function would not lead to the best parameters. Errors when many people remain to alight propagate in the cumulative flow. In addition, processes with low volumes of passengers are the majority (see Fig. 4). Then, we added a factor to penalize the cases at low remaining volume. Therefore, the cost function is a traditional quadratic cost function with the addition of a factor increasing the weight of the term proportional to the number of people remaining to alight (or board, or both).

Then, we minimize the cost function C_X:

$$C_X = \sum_{n=(k,s,d,j,t)}\left(\left(\frac{dX}{dt}\right)_n^{real} - \psi_X\left(1-\exp(-\gamma_X\cdot\sqrt{X_{\infty,n}^{real}-X_n^{real}})\right)\left(\exp(-\gamma_X\cdot\sqrt{X_{\infty,n}^{real}-X_n^{real}})\right)\right)^2\times(X_{\infty,n}^{real}-X_n^{real}) \tag{6}$$

The same cost function is considered to fit parameters for A, B, and N. The minimum is obtained using the function *minimize* from *scipy.optimize* in Python. The fit is sensitive to initial conditions. Then, we did 500 fits where we randomly chose initial conditions and selected the best solution overall. The data used corresponds to different sub-samples presented in Section 2. These data sub-samples are from similar enough setups to fit a single set of parameters for each sub-sample. The parameters then feed the stochastic model.

3.2. Stochastic modeling of the alighting and boarding process

3.2.1. Alighting and boarding process: a Markov process

The deterministic model presented in the previous section aims to describe the average dynamics of the alighting and boarding process. Any passenger movement is yet subject to important variability due to various causes. A stochastic approach is presented in this section to model the variability of time between two alightings, two boardings, or two movements: variability due to individuals (body shape, desired speed...), variability due to interaction between individuals (distance between individuals), variability due to irregular organization (variability of the density). All these variabilities are considered together in a stochastic process: $(X(t))_{t\geq 0}$.

Contrary to the previous subsection, $(X(t))_{t\geq 0}$ is no longer a deterministic variable; it is now a stochastic process. $(X(t))_{t\geq 0}$ is assumed to be a continuous-time Markov process: moving from state $X(t)=i$ to another state only depends on the flow conditions at state $X(t)=i$. The only accessible state from state i is $i+1$, except for X_∞, which is absorbing. We have:

$$P(X(t+h)=j\,|\,X(t)=i) = \delta_{i,j}(1-\mu_{i+1}h) + \delta_{i+1,j}\mu_{i+1}h + o(h), \tag{7}$$

where $\delta_{i,j}$ is the Kronecker symbol (whose value is 1 if $i\neq j$, 0 otherwise), h is a time increment, and μ_{i+1} is the transition rate which monitors how fast the probability of the next movement will increase with time.

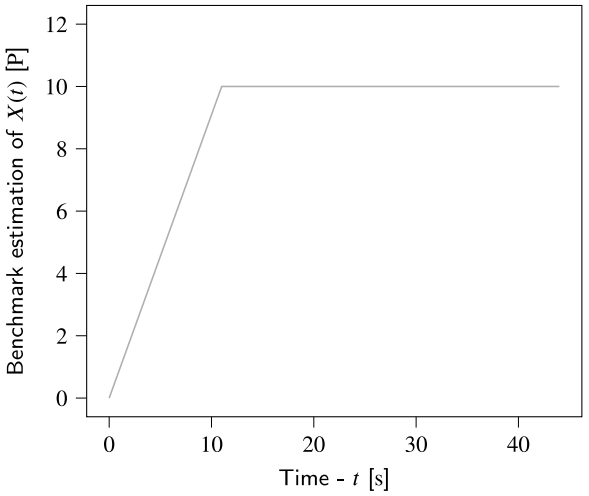

Fig. 6. Illustration of the linear model selected as the benchmark in this work for A, B or N, in the case of $X_\infty = 10$P.

Then, as the process is assumed to be a Markov process, the time separating the ith movement and the $(i+1)^{th}$ movement follows an exponential distribution whose average is the inverse of the transition rate $\left(\frac{1}{\mu_{i+1}}\right)$. Based on the above-derived deterministic model, the transition rate is given by:

$$\frac{1}{\mu_i} = \frac{dt}{dX} = \frac{1}{\psi_X \left(1 - \exp(-\gamma_X \cdot \sqrt{X_\infty - i})\right)\left(\exp(-\gamma_X \cdot \sqrt{X_\infty - i})\right)}, \tag{8}$$

where ψ_X and γ_X are the parameters fitted for the deterministic model. The stochastic model is not a Poisson process because the transition rate μ_i is not assumed to be constant with i a priori. An important observation to be made here is that the above-proposed deterministic model is the average of the proposed stochastic model.

This stochastic model can be seen as a queueing model where arrivals do not follow a Poisson process but a Markov process with varying transition probabilities (due to varying densities), and the service time is infinite.

Such a model enables the simulation of various alighting and boarding processes at the scale of individuals. In practice, this model is particularly interesting for evaluating the alighting and boarding time distribution.

3.2.2. Simulating alighting and boarding time distributions

The proposed stochastic model is used to simulate the alighting and boarding processes from which we compute the alighting and boarding times. The method used to compute the alighting and boarding time is the quantile method developed by Baali et al. [76] using the parameter $q = 0.8$.

From a given number of alighting (A_∞), boarding (B_∞), or total (N_∞), 1,000 alighting and boarding processes are simulated. The resulting distribution of alighting and boarding times is then compared to the distribution obtained from the real-world data (data from the sensors).

3.3. Benchmark model

In the literature, alighting and boarding times are often assumed to be linear with the number of alightings and boardings for practical use [77,78]. This linear model is often taken as a reference for comparison with other models, in particular for buses [46,79]. In this study, we also consider a linear model as a benchmark. A determinist benchmark is as follows:

$$X = \begin{cases} \eta t, & \text{if } \eta t \leq X_\infty \\ X_\infty, & \text{otherwise} \end{cases} \tag{9}$$

where η is fitted the same way as the deterministic model. An illustration is proposed in Fig. 6.

A stochastic benchmark is also considered, which is a queueing model M/D/1 with an infinite service time. Moving from state $X(t) = i$ to another state only depends on state $X(t) = i$, and the only accessible state from state i is $i + 1$, except for X_∞, which is an absorbing state.

$$P(X(t+h) = j | X(t) = i) = \delta_{i,j}(1 - \eta h) + \delta_{i+1,j}\eta h + o(h). \tag{10}$$

where η (the service rate) is the parameter of the deterministic benchmark model. This model is similar to the one proposed by [19]. The same observation can be made here: the deterministic benchmark is the average of the stochastic benchmark.

The stochastic benchmark is also simulated 1,000 times to obtain the alighting and boarding time distribution for fixed A_∞ (or B_∞ or N_∞).

Notice that two-dimensional models, such as cellular automata or social force models, would have been relevant benchmarks. However, given the absence of two-dimensional data sources (for example, videos) to calibrate them, we have not selected them.

3.4. Model assessment

To assess the model and the benchmark, we shuffled the different stops in the data, divided it into five samples, and performed a 5-fold cross-validation (with 4/5 of the data being the train set and 1/5 of the data being the test set).

The quality of the fit is assessed on $X(t)$, across the entire test data sample. The statistical indicators used are the mean absolute error (MAE), the root mean square error (RMSE), and the cost function.

A MAE of 1 $\mathrm{P}\cdot\mathrm{s}^{-1}$ means that, on average among the counting events, the model is 1 $\mathrm{P}\cdot\mathrm{s}^{-1}$ away from the real alighting rate, or boarding rate, or movement rate. The RMSE gives more importance to the counting events where the model is further away from the real data. Values of MAE and RMSE, along with the cost functions, are averaged over the five test samples of the 5-fold cross-validation. A confidence interval at 5% threshold is also computed to assess the significance of the difference between the model and the benchmark fitting.

Both the model and benchmark are compared with the alighting and boarding time distributions from the real data. The Kullback–Leibler divergence [80] between the model (or the benchmark) and real data is used for the comparison. Another assessment is done based on the 2nd and 8th deciles of the distributions because the main interest for practical use is the central part of the distributions, let alone the existence of some outliers in the real-world data.

Concerning the parameters obtained from the fit, their averages over the 5-fold cross-validation are compared to values obtained in the literature. The standard deviation among the five values is used to calculate confidence intervals at a 5% threshold.

If the deterministic model is assessed via a 5-fold cross-validation, it is not the case with the stochastic model. This is because too small amounts of data are available to build distributions for each total number of movements. This leads to correlations between our model's distributions and distributions from real data.

4. Results

4.1. Visual assessment

4.1.1. Visuals of the deterministic models

Fig. 7 shows an example of the profile of cumulative alighting, cumulative boarding for uni-directional flows and cumulative movements for bi-directional flows obtained with the deterministic model (once the differential equation is solved), the deterministic benchmark, and the real data for the different occurrences having the same total number of alighting, boarding, or movements. With a higher volume (see right of the figure), despite an important spread of the real data, the model seems to fit the alighting and boarding process better than the benchmark. With a smaller volume (left of the figure), little difference is observed between the benchmark and the model: both correctly fit the real data. In any case, the benchmark seems to better fit the beginning of the alighting and boarding process. This is probably due to the weighting term in the cost function of the fit, which provides more importance to cases where many people are left to alight and/or board (i.e, the beginning).

4.1.2. Visual of the stochastic model

A similar qualitative analysis can be done for the stochastic model and the stochastic benchmark. For a given cumulative flow X, the 2nd and the 8th deciles of t are plotted for real data, the stochastic model, and the stochastic benchmark (see Fig. 8). The figure only shows the bi-directional case because insufficient data were available for uni-directional cases to get robust deciles. Both the stochastic model and the stochastic benchmark correctly reproduce the spread of the alighting and boarding process. The model shows superiority for a higher total number of movements (left of the figure). The stochastic model fits particularly well the end of the alighting and boarding process in both cases, where the spread increase is more important. In contrast, the stochastic benchmark seems to better fit the beginning of the alighting and boarding process, similar to what was observed in the deterministic case.

4.2. Quality of the fits

Then, we assess the performance of the proposed deterministic model compared to the deterministic benchmark. Table 5 shows these performances with MAE, RMSE, and the cost function, along with the relative confidence intervals. It is not worth assessing the stochastic model and the stochastic benchmark with these averaged indicators because the average of the stochastic model and benchmark corresponds to their deterministic counterparts. The results in this section mainly provide insight into the quality of the fit.

For the uni-directional alighting and the bi-directional case, the MAE and RMSE are lower with the model (with a gain of about 20%). The improvements of the MAE and RMSE with our model in both cases are statistically significant, given the confidence intervals computed using the 5-fold cross-validation. For the uni-directional boarding case, no difference in MAE is observed. This result is consistent with the observations of Kuipers et al. [60], which showed no influence of volume on boarding flow sizes, contrary to alighting flow sizes.

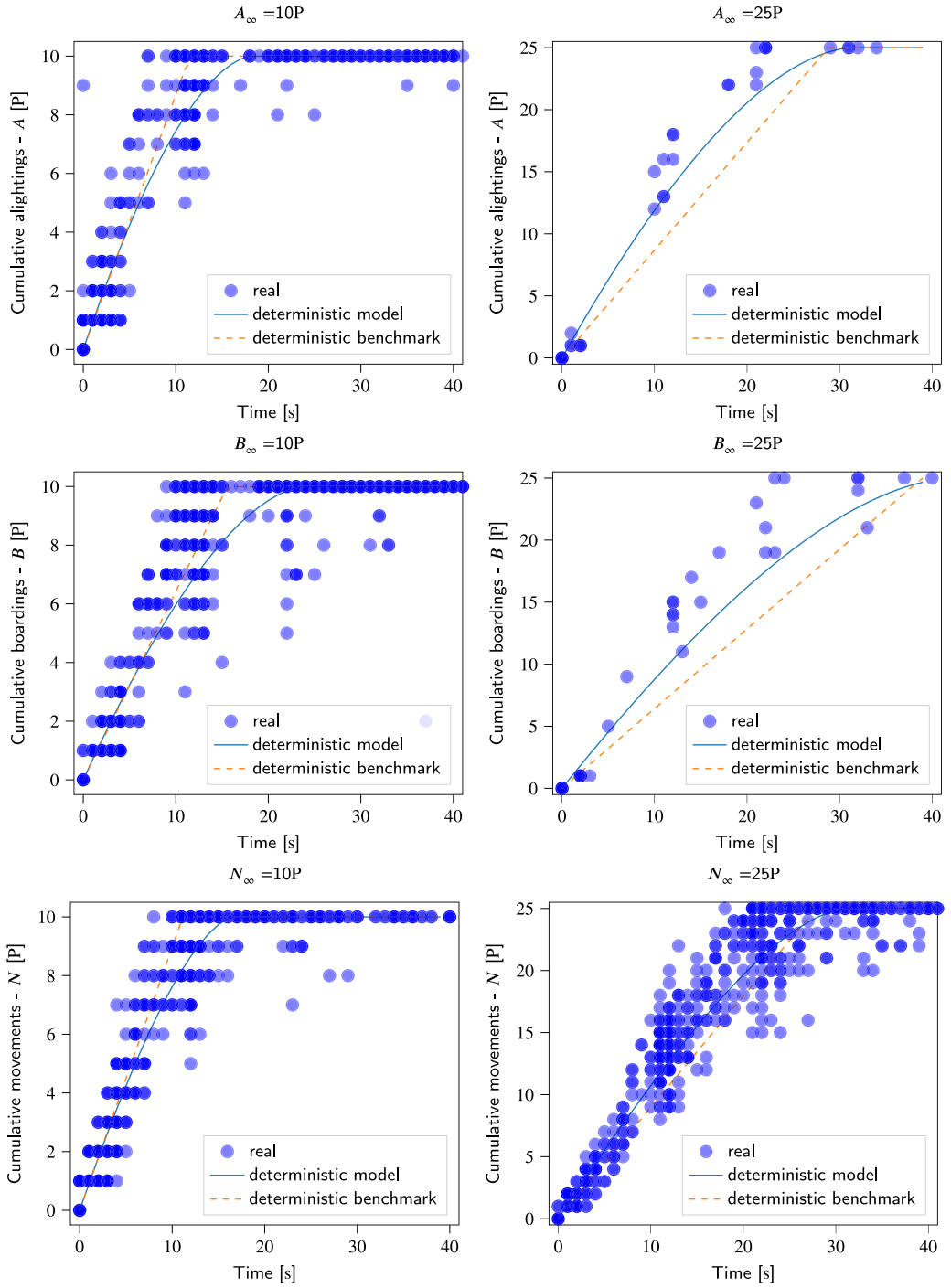

Fig. 7. Examples of the profile of uni-directional cumulative alighting (top), uni-directional cumulative boarding (center), and bi-directional cumulative movements (bottom) from real data, the deterministic model, and the deterministic benchmark for processes having a total of 10 movements (left) and a total of 25 movements (right). The superiority of the model is visible in higher total volumes.

However, for all cases, the RMSE is much better for the model. This shows that larger errors are lower with the model. As the error is assessed on the derivative of A, B, and N, lower cumulative errors will be observed with the model. This advocates for the superiority of the model compared to the benchmark. Notice that the RMSE for the bi-directional case is higher than in both uni-directional cases; this can be due to hard-to-predict hindering of both directions of the flows that we have chosen to neglect.

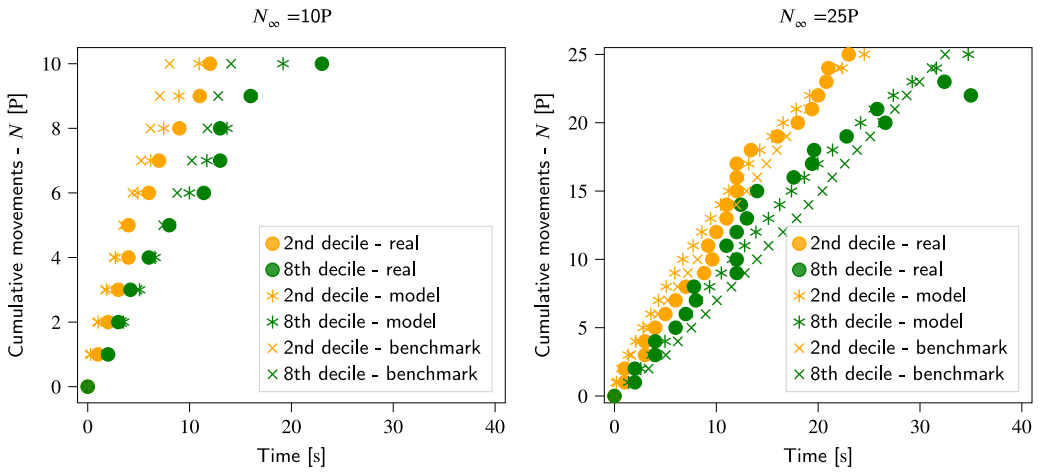

Fig. 8. Examples of the 2nd and 8th deciles of times in the dynamic of bi-directional cumulative movements from real data, the stochastic model, and the stochastic benchmark for processes having a total of 10 movements (left) and a total of 25 movements (right). The superiority of the model is visible for higher total volumes and at the end of the alighting and boarding process.

Table 5
Statistical performances of the fits of the deterministic model compared to the deterministic benchmark. The model does better in MAE except for uni-directional boarding and better in RMSE for all cases. The cost function is higher with the model because the cost function penalizes more at the beginning of the alighting process.

Case	MAE [P · s^{-1}]	RMSE [P · s^{-1}]	Cost function - C_X [P^3·s^{-2}]
Only alighting - Model	0.08 ± 0.002	0.16 ± 0.005	5.54 ± 0.23
Only alighting - Benchmark	0.10 ± 0.004	0.22 ± 0.006	4.98 ± 0.10
Only boarding - Model	0.11 ± 0.001	0.18 ± 0.002	2.61 ± 0.10
Only boarding - Benchmark	0.11 ± 0.002	0.21 ± 0.003	2.37 ± 0.04
Bi-directional - Model	0.10 ± 0.001	0.20 ± 0.002	6.92 ± 0.02
Bi-directional - Benchmark	0.12 ± 0.001	0.24 ± 0.002	6.15 ± 0.01

In any case, the narrow width of the confidence interval computed thanks to the 5-fold cross-validation shows a low dependency of the fit quality on the data. In particular, the variance in the data quality does not influence the overall goodness of the fit.

The cost function is lower for the benchmark, which means that the benchmark better fits the beginning of the alighting and boarding process at high volume. This is consistent with observations made in the visual assessment. This challenges one of the main hypotheses of our model, which is that the alighting and boarding process is only driven by macroscopic density effects. Yet, at the beginning of the process, a discontinuity is generated by the door opening. This is especially the case for only alighting or only boarding, where no one is standing on the other side of the door. The first alighting or boarding movement is not hindered by other passengers, resulting in faster flows. In addition, push-from-the-back effects should further increase the flow at the beginning in such cases.

The next section is dedicated to the assessment of the obtained alighting and boarding time distributions.

4.3. Alighting and boarding time distributions from the stochastic model

In this section, we compare the distribution of alighting and boarding times from real data to the simulated distribution obtained from the stochastic approach with both the benchmark and the model. Each distribution is computed for a fixed total number of movements, and the boarding/alighting ratio is neglected. We discuss only cases with a minimum of 30 values in the distribution. Hence, we focus on the bi-directional case, which is much more frequent in the dataset.

As a first illustration, Fig. 9 shows the distributions of alighting and boarding times for a total number of movements of 10P and 25P from real data, the stochastic model, and the stochastic benchmark. For a total of 10 movements, the stochastic model and the stochastic benchmark have a similar distribution, both close to real data. For a total of 25 movements, the stochastic model remains close to real data, while the stochastic benchmark has a significantly different distribution. In both cases, a non-negligible number of outliers is observed in real data, whereas this is neither the case for the stochastic model nor for the stochastic benchmark. We have no insight into whether these outliers illustrate real-world situations or come from data issues, but the goal of our approach is to have a decent prediction for the central part of the distribution, representing the majority of real-world situations. Therefore,

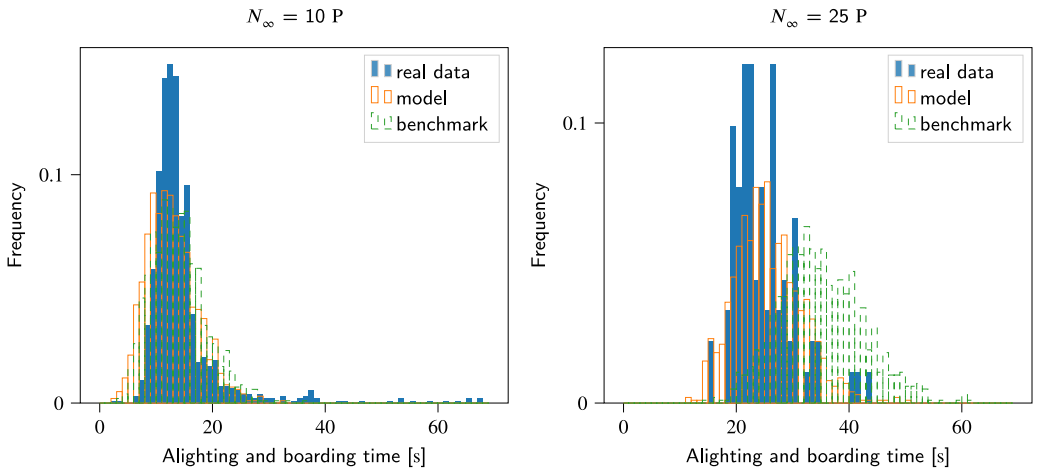

Fig. 9. Distributions of alighting and boarding times for a total number of movements of 10P and 25P from real data, the stochastic model, and the stochastic benchmark. The stochastic model decently fits the distribution of alighting and boarding times from real data in both cases, while the stochastic benchmark only does so for 10 movements. Several outliers are observed in real data.

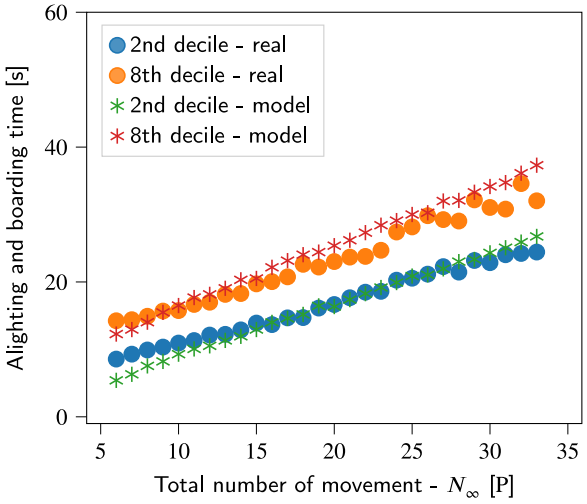

Fig. 10. Evolution of the 2nd and 8th deciles in alighting and boarding times from real data (asterisks) and from the stochastic model (dots). Deciles from real data are well predicted by deciles of the model.

these outliers fall out of the scope of this analysis. This is why we focus on the 2nd and 8th deciles of the alighting and boarding times.

Fig. 10 shows the evolution of the 2nd and 8th deciles of the alighting and boarding times distributions from real data and the stochastic model. From this decile analysis, the model correctly fits the real data, especially for higher volumes. In particular, the model correctly captures the increase in the interval between the deciles with increasing volume. The variability increases more slowly than the total number of movements, which is consistent with observations made by Buchmueller et al. [40].

Fig. 11 shows the Kullback–Leibler divergence in alighting and boarding times distributions between the model (respectively, the benchmark) and real data. The Kullback–Leibler divergence is lower with the model above 23 movements overall and below 12 movements overall, while being equivalent to the benchmark elsewhere. The lowest Kullback–Leibler divergences between the stochastic benchmark and real data are obtained for intermediate values (between 12 and 19 passengers). This is probably due to the cost function's form, which penalizes processes with few people overall (from the weighting factor), and processes with few observations to fit the data. Concerning the stochastic model, the Kullback–Leibler divergence decreases with the total number of movements. This highlights that the model's predictive power is slightly better at higher volumes, consistent with observations made in Fig. 8.

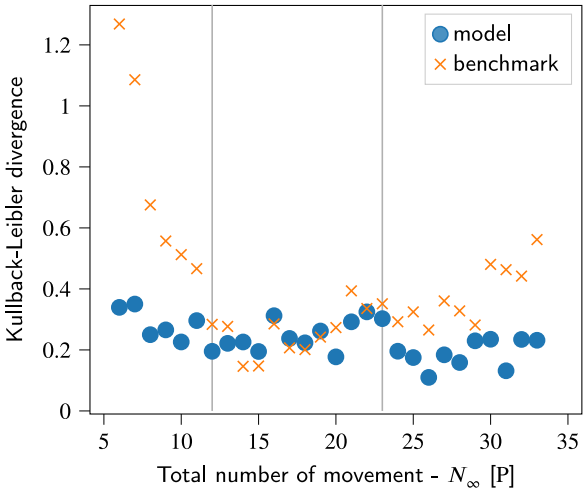

Fig. 11. Kullback–Leibler divergence for the alighting and boarding times distributions between the stochastic model (in blue dots) or the stochastic benchmark (in orange crosses) and real data. The distributions with the model are closer to real data, with a higher total number of passengers.

4.4. Physical interpretation of the parameters

The parameters obtained when fitting the model have physical meanings. As a reminder, $\psi = v_0 D \rho_{max}$, where v_0 is the free-flow speed, ρ_{max} is the maximal density, D is the door width (fixed). ρ_{max} could vary according to passenger body shapes (body size, body weight, luggage, etc.). v_0 depends on people's behavior. With fixed γ, a higher value of ψ means a higher flow for a similar number of alighting, boarding, or movements; and, in particular, a higher maximal flow $\left(\frac{dX}{dt}\right)_{max}$. From the value of γ, it is possible to derive the critical demand X_{crit} (number of alighting, boarding, or movement corresponding to the highest flow). The critical demand can be obtained by deriving the left-hand side of Eq. (5) with respect to the variable $X_\infty - X$.

4.4.1. All data

Table 6 presents the parameters obtained with their confidence intervals. Values for bi-directional flows are similar to those of uni-directional boarding flows with the same confidence intervals. The flow size (ψ) for alighting flows is higher with wider confidence intervals, which means more variability when there is only alighting. This is probably due to variability in the configuration of passengers inside the train before alighting (people in front of the door versus people seated further in the carriage), which should influence the alighting rate for a fixed number of passengers.

Though a higher value for the uni-directional alighting case, the width of the confidence interval remains low compared to the parameter. This shows that the variance in the data quality does not influence the fitted parameters.

For similar volumes, the uni-directional alighting rate is significantly higher than the uni-directional boarding rate. This result is consistent with the literature. For example, Fu et al. [64] also found significantly higher flows for only alighting than for only boarding (in the case of a bi-directional flow with full compliance to the alighting priority). However, they found alighting rates closer to boarding rates in generic bi-directional flows as shown by Kuipers et al. [60] or Seriani et al. [41]. This result is not reproducible with our approach as we amalgamate alighting and boarding in generic bi-directional flows. Similarly to how we amalgamate alighting and boarding in bi-directional flows, Fu et al. [64] found movement rates closer to uni-directional boarding flows compared to uni-directional alighting flows, which is consistent with our results.

The critical demand that we obtain should be considered with caution. Fig. 4 showed that no situation having volumes higher than the critical demand was observed in the data. This means that these critical demands are projections from low-volume situations. This also means that, based on our results, no congested regime is observed in the data. The critical demand is particularly high for uni-directional flows (alighting or boarding). Knowing the capacity of one car of the train is 132P, the model predicts that no congested regime may occur in any only alighting or only boarding situations. In the case of a bi-directional flow, it would be possible to see a congested regime, but data of higher demand are necessary for further discussion.

In Table 7, the values of the obtained parameters ψ are compared with those obtained from comparable flow-density relationships in the pedestrian literature (i.e., Greenshields formulations). For only alighting, ψ is close to the values of level walkways in the literature, where ψ is generally higher. For only boarding or bi-directional flows, it is close to the values for ascending stairs. The difference between uni-directional and bi-directional cases in the literature is generally low, and there is no trend on which type of flow should have a higher ψ than the other. This is consistent with the similarity observed in only boarding flows and bi-directional flows from our model. In further detail, uni-directional and bi-directional flows have different free-flow speed and maximal density parameters in the literature; these parameters cannot be explicitly obtained with our model.

Table 6
Parameters obtained from the three fits done and their confidence intervals at 5% threshold. Parameters from the bi-directional and boarding flow are similar with similar confidence intervals. ψ is much higher for the alighting flow with a much larger confidence interval.

Case	$\psi = Dv_0\rho_{max}$ [P · s^{-1}]	γ	$\left(\frac{dX}{dt}\right)_{max}$ [P · s^{-1}]	X_{crit} [P]
Only alighting	9.1 ± 0.66	0.04 ± 0.003	2.28	300
Only boarding	5.0 ± 0.04	0.06 ± 0.001	1.25	133
Bi-directional	4.9 ± 0.04	0.09 ± 0.001	1.23	56

Table 7
Comparison of the parameters ψ from our model and comparable existing works (with Greenshields formulation). The parameter obtained for uni-directional alighting flow is close to the case of level walkways. Cases of uni-directional boarding or bi-directional flows are more in line with cases of ascending stairs.

Case	Context	v_0 [m · s^{-1}]	ρ_{max} [P · m^{-2}]	$\psi = Dv_0\rho_{max}$ [P · s^{-1}]
Only alighting (our model)	Train door			9.1 ± 0.66
Only boarding (our model)	Train door			5.0 ± 0.04
Bi-directional (our model)	Train door			4.9 ± 0.04
Fruin [69]	Level uni-directional	1.43	4.09	9.20
Fruin [69]	Level bi-directional	1.36	4	8.70
Fruin [69]	Ascending stairs	0.57	7.40	6.75
Fruin [69]	Descending stairs	0.65	6.70	6.97
Nelson [82]	Level	1.4	3.8	8.51
Nelson [82]	Any stairs	0.8	3.8	4.86
Ye et al. [83]	Level uni-directional	1.44	3.2	7.38
Ye et al. [83]	Level bi-directional	1.19	4.11	7.83
Ye et al. [83]	Descending stairs	0.97	6.06	9.41
Ye et al. [83]	Ascending stairs	0.69	4.93	5.44

Existing studies show a higher maximal flow in the fundamental diagram for pedestrian crossing bottlenecks compared to other walkways' results [54,81]. This would lead to a higher ψ, which is not observed here. The bottlenecks studied in the work of Seyfried et al. [54] were narrower than the 1.6 m door studied here. This might be a potential explanation.

4.4.2. Specific sub-samples

For bi-directional flows (see Table 8), we compare the obtained parameters varying non-compliance, day-period, and for a specific station. A first observation is that the values are close for the two parameters in all cases.

Compared to peak hours, ψ in off-peak hours is slightly lower, probably meaning a lower v_0. Bosina and Weidmann [84] showed a slight but not significant increase in walking speed in peak hours, in line with our result. γ is similar in both cases, which means no influence of peak hours on the relationship between volumes and density, and an increase in flow for similar volumes. Other works, such as the study of Wiggenraad [85], showed a very low increase in alighting and boarding rates during peak hours, similar to here.

Regarding non-compliance, higher non-compliance has a not significantly higher ψ while having a lower γ. A lower γ means that the density increases more slowly with the remaining number of passengers to alight or board. This contradicts findings from Baali et al. [68], where higher densities were found with higher non-compliance and the same volumes. The confidence interval of both parameters at higher non-compliance is wider than in any other case. This can be due to a more variable flow at higher non-compliance between one situation and another. For example. Yang et al. [33] showed more fluctuation in the alighting and boarding dynamics for non-compliance cases from simulation results.

For *Versailles-Chantiers*, the parameters are very similar to the case with all the data. This confirms that stations have negligible to no influence on the modeling of the alighting and boarding process. *Versailles-Chantiers* being, moreover, the station with the highest demand in *Line N*, this confirms one major hypothesis of the approach that the model considers intrinsically the volume with parameters independent of it.

5. Conclusion

In this work, we have presented an innovative approach to model the alighting and boarding process. This approach differs from what is generally done by discussing the alighting and boarding process as a cumulative flow problem rather than a two-dimensional pedestrian movement. Two models have been discussed: a deterministic model depending on alighting and boarding volumes fitted on disaggregated passenger counting data; and a stochastic model, which is a Markov process whose average is the deterministic

Table 8
Comparison of the parameters for bi-directional flows between peak hours and off-peak hours; lower non-compliance and higher non-compliance; at the station of *Versailles-Chantiers*. $\psi = v_0 D \rho_{max}$.

	ψ_N [P s^{-1}]	γ_N [$P^{-\frac{1}{2}}$]
All	4.9 ± 0.04	0.09 ± 0.001
Peak hours	4.9 ± 0.02	0.09 ± 0.0004
Off-peak hours	4.8 ± 0.05	0.09 ± 0.002
Lower non-compliance	5.0 ± 0.03	0.09 ± 0.001
Higher non-compliance	5.1 ± 0.16	0.07 ± 0.003
Versailles-Chantiers	4.8 ± 0.04	0.09 ± 0.001

model. The disaggregated passenger counting data used are particularly suitable for calibrating the model. The deterministic model has then been compared with a linear benchmark and has shown better predictive power for uni-directional alighting flows and bi-directional flows (with a gain of around 20% in MAE and RMSE). The stochastic model helps capture the variability of the flow, and the probability distribution of alighting and boarding times for a given volume is correctly reproduced. The physical parameters obtained from the fit are in line with existing literature. In summary, the proposed approach is a data-driven stochastic model that precisely fits time-dependent alighting and boarding rates, and its parameters are physically consistent.

This novel approach is a relevant addition to existing models for the alighting and boarding process, provided disaggregated passenger counting data is available. The proposed approach is at the frontier between data-driven approaches (where the physical meanings of the parameters are rarely interpretable) and microscopic pedestrian simulations (where it is difficult to calibrate the input parameters). Our model completes the gap in the spectrum between these two types of approaches.

One can suggest dwell time planning as a practical implication. Forecasting alighting and boarding time distributions at the door level would help assess train dwell times. Train dwell times can be computed taking the maximum of the difference between the scheduled departure time and the observed arrival time, and the maximum alighting and boarding time among the train doors, with the addition of the technical time (necessary time to open/close the doors and dispatch). The proposed model provides an estimate of the alighting and boarding time for each door. This would enable building more robust schedules to reduce dwell time delays. Based on the model, if the maximum alighting and boarding time among the doors (plus the technical time) is regularly larger than the scheduled dwell time, dwell time delays will occur on a regular basis. Similar analyses have been performed by Coulaud et al. [44] with a data-driven prediction or Tortainchai et al. [86] using a data envelopment analysis. In the case of our work, once calibrated on a rolling stock, the model can be used to schedule dwell times even without disaggregated counting data.

The model proposed in this article may also be integrated into a mesoscopic railway simulator for a line-scale or network-scale simulation. Dwell time models were integrated into simulators by D'Acierno et al. [77] or Zhang et al. [63], for example, but they did not describe the entire alighting and boarding process. An agent-based pedestrian simulator for passengers through a railway network has also been developed by Hänseler et al. [87].

Another implication of our model is the analysis of the efficiency of dwell time (necessary dwell time versus number of alightings and boardings) based on total volumes of passengers and their distribution among the doors. Unlike existing works [86,88], the efficiency of dwell time can be obtained analytically given the number of passengers alighting and boarding at each door. Such a dwell time efficiency analysis can be used to assess the relevance of a given rolling stock to run on a new line.

This model is built to deal with cumulative flows. Then, it is complicated to compare its results with two-dimensional pedestrian models from the literature, especially because no two-dimensional data (video) is available to calibrate two-dimensional models. Our model would benefit from a comparison to more sophisticated existing approaches, such as cellular automata or social force models. In addition, the data used for this work gave information on alightings and boardings every 3s on average. Such gaps between two counting events make the alighting/boarding rates rather imprecise, and obtaining a good fit is made more difficult.

In future work, we will extend the approach to other lines with higher demands, which will be possible with the recent arrival of new rolling stocks on the Paris RER. In this context, it will be crucial to consider on-board and on-platform people who neither alight nor board. Improvements of the model will be necessary, including a better modeling of bi-directional flows, with the addition of stochastic choices of whether the following movement is an alighting or a boarding. This addition will help consider different rates for alighting and boarding in bi-directional flows, and a direct consideration of the boarding/alighting ratio that has been neglected so far, and whose influence will be significant for higher volumes. Other functions for the fundamental diagram and the density function will also have to be assessed to deal with higher congestion levels and to challenge the results presented here. The behavior at the beginning of the alighting and boarding process regarding discontinuity after the door opening will also be studied for better modeling. Finally, this approach will be compared to two-dimensional alternatives.

CRediT authorship contribution statement

Mehdi Baali: Writing – original draft, Validation, Methodology, Conceptualization. **Christine Buisson:** Writing – review & editing, Supervision, Conceptualization. **Rémi Coulaud:** Writing – review & editing, Supervision, Data curation. **Winnie Daamen:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mehdi Baali reports financial support was provided by ANRT. Mehdi Baali reports a relationship with SNCF Voyageurs that includes: employment and non-financial support. Remi Coulaud reports a relationship with SNCF Voyageurs that includes: employment. Christine Buisson reports a relationship with ENTPE Civil Engineering and Building Education and Research Department that includes: employment. Winnie Daamen reports a relationship with Delft University of Technology that includes: employment. Mehdi Baali reports a relationship with Gustave Eiffel University that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Computation of the critical demand and the maximal flow

Critical demand The critical demand is the number of remaining movements (alighting and/or boarding) $X_\infty - X$ for which the flow is maximal. To compute it, we use the left-hand side of the equation:

$$\frac{dX}{dt} = \psi \left(1 - \exp(-\gamma \cdot \sqrt{X_\infty - X(t)})\right) \left(\exp(-\gamma \cdot \sqrt{X_\infty - X(t)})\right).$$

We compute the derivative of:

$$g(y) = \psi \left(1 - \exp(-\gamma \cdot \sqrt{y})\right) \left(\exp(-\gamma \cdot \sqrt{y})\right),$$

which is:

$$\frac{dg}{dy} = \frac{\psi\gamma}{2\sqrt{y}} \left(2\exp(-2\gamma \cdot \sqrt{y}) - \exp(-\gamma \cdot \sqrt{y})\right).$$

This derivative is zero for:

$$y = X_{crit} = \left(\frac{\log(2)}{\gamma}\right)^2,$$

which gives the critical demand.

Maximal flow The maximal flow is the flow at the critical demand, but it can be directly obtained from Eq. (5). Writing $\lambda_X = \exp(-\gamma \cdot \sqrt{X_\infty - X(t)})$, we have:

$$\frac{dX}{dt} = \psi \left(1 - \lambda_X\right) \lambda_X.$$

From there, we can get the maximal flow:

$$\left(\frac{dX}{dt}\right)_{max} = \psi \times 0.25.$$

Data availability

The data that has been used is confidential.

References

[1] R. Kuipers, C.-W. Palmqvist, Passenger volumes and dwell times for commuter trains: A case study using automatic passenger count data in stockholm, Appl. Sci. 12 (2022) http://dx.doi.org/10.3390/app12125983.
[2] Anupriya, D.J. Graham, P. Bansal, D. Hörcher, R. Anderson, Optimal congestion control strategies for near-capacity urban metros: Informing intervention via fundamental diagrams, Phys. A 609 (2023) 128390.
[3] T. Seo, K. Wada, D. Fukuda, A macroscopic and dynamic model of urban rail transit with delay and congestion, in: Transportation Research Board 96th Annual Meeting, 2017.
[4] D. Yazdani, M.N. Omidvar, I. Deplano, C. Lersteau, A. Makki, J. Wang, T.T. Nguyen, Real-time seat allocation for minimizing boarding/alighting time and improving quality of service and safety for passengers, Transp. Res. Part C: Emerg. Technol. 103 (2019) 158–173.
[5] D.C. Duives, W. Daamen, S.P. Hoogendoorn, State-of-the-art crowd motion simulation models, Transp. Res. Part C: Emerg. Technol. 37 (2013) 193–209.
[6] H. Dong, M. Zhou, Q. Wang, X. Yang, F.-Y. Wang, State-of-the-art pedestrian and evacuation dynamics, IEEE Trans. Intell. Transp. Syst. 21 (5) (2019) 1849–1866.
[7] L. Henderson, The statistics of crowd fluids, Nature 229 (5284) (1971) 381–383.
[8] R.L. Hughes, A continuum theory for the flow of pedestrians, Transp. Res. Part B: Methodol. 36 (6) (2002) 507–535.
[9] V.J. Blue, J.L. Adler, Emergent fundamental pedestrian flows from cellular automata microsimulation, Transp. Res. Rec. 1644 (1) (1998) 29–36.
[10] A. Schadschneider, A. Kirchner, K. Nishinari, CA approach to collective phenomena in pedestrian dynamics, in: International Conference on Cellular Automata, Springer, 2002, pp. 239–248.
[11] D. Helbing, P. Molnar, Social force model for pedestrian dynamics, Phys. Rev. E 51 (5) (1995) 4282.
[12] H. Xi, S. Lee, Y.-J. Son, An integrated pedestrian behavior model based on extended decision field theory and social force model, Human-in-the-Loop Simulations: Methods Pr. (2011) 69–95.

[13] S. Paris, J. Pettré, S. Donikian, Pedestrian reactive navigation for crowd simulation: a predictive approach, in: Computer Graphics Forum, Vol. 26, Wiley Online Library, 2007, pp. 665–674.
[14] M. Moussaïd, D. Helbing, G. Theraulaz, How simple rules determine pedestrian behavior and crowd disasters, Proc. Natl. Acad. Sci. 108 (17) (2011) 6884–6888.
[15] S.P. Hoogendoorn, P.H. Bovy, Pedestrian route-choice and activity scheduling theory and models, Transp. Res. Part B: Methodol. 38 (2) (2004) 169–190.
[16] M. Xiong, W. Cai, S. Zhou, M.Y.H. Low, F. Tian, D. Chen, D.W.S. Ong, B.D. Hamilton, A case study of multi-resolution modeling for crowd simulation, in: Proceedings of the 2009 Spring Simulation Multiconference, 2009, pp. 1–8.
[17] T. Robin, G. Antonini, M. Bierlaire, J. Cruz, Specification, estimation and validation of a pedestrian walking behavior model, Transp. Res. Part B: Methodol. 43 (1) (2009) 36–56.
[18] F.E.H. Wijermans, Understanding crowd behaviour: simulating situated individuals, 2011.
[19] G.G. Løvås, Modeling and simulation of pedestrian traffic flow, Transp. Res. Part B: Methodol. 28 (6) (1994) 429–443.
[20] W. Daamen, SimPed: a pedestrian simulation tool for large pedestrian areas, in: Congres EuroSIW, Simulation Interoperability Standards Organization, 2002, pp. 1–11.
[21] F. Martinez-Gil, M. Lozano, F. Fernández, MARL-Ped: A multi-agent reinforcement learning based framework to simulate pedestrian groups, Simul. Model. Pr. Theory 47 (2014) 259–275.
[22] X. Song, D. Han, J. Sun, Z. Zhang, A data-driven neural network approach to simulate pedestrian movement, Phys. A 509 (2018) 827–844.
[23] S. Hoogendoorn, W. Daamen, Design assessment of lisbon transfer stations using microscopic pedestrian simulation, WIT Trans. the Built Environ. 74 (2004).
[24] W. Daamen, Modelling Passenger Flows in Public Transport Facilities. (Ph.D. thesis), 2002.
[25] W. Daamen, S. Hoogendoorn, Pedestrian traffic flow operations on a platform: observations and comparison with simulation tool SimPed, WIT Trans. the Built Environ. 74 (2004).
[26] W. Scherr, P. Manser, P. Bützberger, Simba mobi: Microscopic mobility simulation for corporate planning, Transp. Res. Procedia 49 (2020) 30–43.
[27] T. Salzmann, B. Ivanovic, P. Chakravarty, M. Pavone, Trajectron++: Dynamically-feasible trajectory forecasting with heterogeneous data, in: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII 16, Springer, 2020, pp. 683–700.
[28] Y. Yuan, X. Weng, Y. Ou, K.M. Kitani, Agentformer: Agent-aware transformers for socio-temporal multi-agent forecasting, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 9813–9823.
[29] T. Gu, G. Chen, J. Li, C. Lin, Y. Rao, J. Zhou, J. Lu, Stochastic trajectory prediction via motion indeterminacy diffusion, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 17113–17122.
[30] M. Dubroca-Voisin, B. Kabalan, F. Leurent, On pedestrian traffic management in railway stations: simulation needs and model assessment, Transp. Res. Procedia 37 (2019) 3–10.
[31] C. Rudloff, D. Bauer, T. Matyus, S. Seer, Mind the gap: Boarding and alighting processes using the social force paradigm calibrated on experimental data, in: 2011 14th International IEEE Conference on Intelligent Transportation Systems, ITSC, IEEE, 2011, pp. 353–358.
[32] J. Yu, Y. Ji, L. Gao, Q. Gao, Optimization of metro passenger organizing of alighting and boarding processes: Simulated evidence from the metro station in Nanjing, China, Sustainability 11 (13) (2019) 3682.
[33] X. Yang, Q. Wang, T. Liu, A fuzzy-theory-based social force model for studying the impact of alighting area width on alighting and boarding behaviors, IEEE Access 8 (2020) 28595–28606.
[34] B. Su, P. Andelfinger, J. Kwak, D. Eckhoff, H. Cornet, G. Marinkovic, W. Cai, A. Knoll, A passenger model for simulating boarding and alighting in spatially confined transportation scenarios, J. Comput. Sci. 45 (2020) 101173.
[35] X. Yang, X. Yang, F. Pan, Y. Kang, J. Zhang, The effect of passenger attributes on alighting and boarding efficiency based on social force model, Phys. A 565 (2021) 125566.
[36] Z. Li, S. Lo, J. Ma, X. Luo, A study on passengers' alighting and boarding process at metro platform by computer simulation, Transp. Res. Part A: Policy Pr. 132 (2020) 840–854, http://dx.doi.org/10.1016/j.tra.2019.12.017.
[37] Q. Zhang, B. Han, D. Li, Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations, Transp. Res. Part C: Emerg. Technol. 16 (5) (2008) 635–649, http://dx.doi.org/10.1016/j.trc.2007.12.001.
[38] L. Sun, G. Yuan, L. Yao, L. Cui, D. Kong, Study on strategies for alighting and boarding in subway stations, Phys. A 583 (2021) 126302.
[39] Y. Qu, X. Liu, J. Wu, Y. Wei, Modeling pedestrian behaviors of boarding and alighting dynamics in urban railway stations, Transp. A: Transp. Sci. 19 (1) (2023) 2035845.
[40] S. Buchmueller, U. Weidmann, A. Nash, Development of a dwell time calculation model for timetable planning, WIT Trans. the Built Environ. 103 (2008) 525–534, http://dx.doi.org/10.2495/CR080511.
[41] S. Seriani, R. Fernandez, N. Luangboriboon, T. Fujiyama, Exploring the effect of boarding and alighting ratio on passengers' behaviour at metro stations by laboratory experiments, J. Adv. Transp. 2019 (2019) 1–12, http://dx.doi.org/10.1155/2019/6530897.
[42] M. Haghani, M. Sarvi, Crowd behaviour and motion: Empirical methods, Transp. Res. Part B: Methodol. 107 (2018) 253–294.
[43] S. Cornet, C. Buisson, F. Ramond, P. Bouvarel, J. Rodriguez, Methods for quantitative assessment of passenger flow influence on train dwell time in dense traffic areas, Transp. Res. Part C: Emerg. Technol. 106 (2019) 345–359, http://dx.doi.org/10.1016/j.trc.2019.05.008.
[44] R. Coulaud, C. Keribin, G. Stoltz, Modeling dwell time in a data-rich railway environment: with operations and passenger flows data, Transp. Res. Part C: Emerg. Technol. 146 (2023) 103980, http://dx.doi.org/10.1016/j.trc.2022.103980.
[45] W. Daamen, S.P. Hoogendoorn, P.H. Bovy, First-order pedestrian traffic flow theory, Transp. Res. Rec. 1934 (1) (2005) 43–52.
[46] L. Sun, A. Tirachini, K.W. Axhausen, A. Erath, D.-H. Lee, Models of bus boarding and alighting dynamics, Transp. Res. Part A: Policy Pr. 69 (2014) 447–460.
[47] B. Hankin, R.A. Wright, Passenger flow in subways, J. Oper. Res. Soc. 9 (2) (1958) 81–88.
[48] B.D. Greenshields, J.R. Bibbins, W. Channing, H.H. Miller, A study of traffic capacity, in: Highway Research Board Proceedings, Vol. 14, Washington, DC, 1935, pp. 448–477.
[49] U. Weidmann, Transporttechnik der fußgänger: transporttechnische eigenschaften des fußgängerverkehrs, literaturauswertung, IVT Schriftenreihe 90 (1993).
[50] M.J. Hurley, D.T. Gottuk, J.R. Hall Jr., K. Harada, E.D. Kuligowski, M. Puchovsky, J.M. Watts Jr., C.J. Wieczorek, et al., SFPE Handbook of Fire Protection Engineering, Springer, 2015.
[51] L.D. Vanumu, K. Ramachandra Rao, G. Tiwari, Fundamental diagrams of pedestrian flow characteristics: A review, Eur. Transp. Res. Rev. 9 (4) (2017) 49, http://dx.doi.org/10.1007/s12544-017-0264-6.
[52] P. Daly, F. McGrath, T. Annesley, Pedestrian speed/flow relationships for underground stations, Traffic Eng. Control. 32 (2) (1991) 75–78.
[53] N. Harris, Increased realism in modelling public transport services, in: Transportation Planning Methods: Volume 11. Proceedings of Seminar H Held at the 22nd PTRC European Transport Forum, University of Warwick, England September 12-16, 1994. Volume P380, 1994.
[54] A. Seyfried, O. Passon, B. Steffen, M. Boltes, T. Rupprecht, W. Klingsch, New insights into pedestrian flow through bottlenecks, Transp. Sci. 43 (3) (2009) 395–406.
[55] R. Fernández, P. Zegers, G. Weber, N. Tyler, Influence of platform height, door width, and fare collection on bus dwell time: laboratory evidence for Santiago de Chile, Transp. Res. Rec. 2143 (1) (2010) 59–66.

[56] R. Thoreau, C. Holloway, G. Bansal, K. Gharatya, T.-R. Roan, N. Tyler, Train design features affecting boarding and alighting of passengers, J. Adv. Transp. 50 (8) (2016) 2077–2088.
[57] Z. Christoforou, E. Chandakas, I. Kaparias, Investigating the impact of dwell time on the reliability of urban light rail operations, Urban Rail Transit 6 (2) (2020) 116–131.
[58] W. Daamen, Y.-c. Lee, P. Wiggenraad, Boarding and alighting experiments: Overview of setup and performance and some preliminary results, Transp. Res. Rec.: J. the Transportation Res. Board 2042 (1) (2008) 71–81, http://dx.doi.org/10.3141/2042-08.
[59] A. Tirachini, Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers, Transp. A: Transp. Sci. 9 (1) (2013) 28–49.
[60] R.A. Kuipers, F. Carlvik, J. Rahm, C.-W. Palmqvist, A naturalistic study into the flow of alighting and boarding passengers of commuter trains, Transp. Res. Interdiscip. Perspect. 31 (2025) 101369.
[61] S. Seriani, R. Fernandez, Pedestrian traffic management of boarding and alighting in metro stations, Transp. Res. Part C: Emerg. Technol. 53 (2015) 76–92, http://dx.doi.org/10.1016/j.trc.2015.02.003.
[62] N.G. Harris, F. De Simone, B. Condry, A comprehensive analysis of passenger alighting and boarding rates, Urban Rail Transit 8 (1) (2022) 67–98.
[63] J. Zhang, A. Wu, W. An, L. Hu, J. Zhu, State-dependent multi-agent discrete event simulation for urban rail transit passenger flow, Phys. A 652 (2024) 130031.
[64] L. Fu, Q. Chen, Q. Shi, Y. Chen, Y. Shi, Characteristics of pedestrians' alighting and boarding process in metro stations, Tunn. Undergr. Space Technol. 141 (2023) 105362.
[65] A.M. Wahaballa, M. Abdelhaleem, K. Saeed, A. Othman, Experimental analysis of boarding and alighting behavior in urban public transport network: A case study, Transp. Res. Procedia 62 (2022) 25–34.
[66] S. Seriani, J.M. Barriga, A. Peña, A. Valencia, V. Aprigliano, L. Jorquera, H. Pinto, M. Valenzuela, T. Fujiyama, Analyzing the effect of crowds on passenger behavior inside urban trains through laboratory experiments — A pilot study, Sustainability 14 (22) (2022) 14882, http://dx.doi.org/10.3390/su142214882.
[67] Y. Qu, Y. Xiao, H. Liu, H. Yin, J. Wu, Q. Qu, D. Li, T. Tang, Analyzing crowd dynamic characteristics of boarding and alighting process in urban metro stations, Phys. A 526 (2019) 121075.
[68] M. Baali, C. Buisson, R. Coulaud, M. Moussaïd, The impact of passenger non-compliance during alighting and boarding on railway stations..
[69] J.J. Fruin, Pedestrian Planning and Design, Technical Report, 1971.
[70] J.L. Pauls, Building evacuation: research findings and recommendations, Fires Hum. Behav. (1980) 251–275.
[71] W.H. Lam, J.F. Morrall, H. Ho, Pedestrian flow characteristics in Hong Kong, Transp. Res. Rec. 1487 (1995) 56–62.
[72] R. Rastogi, S. Chandra, et al., Pedestrian flow characteristics for different pedestrian facilities and situations, 2013.
[73] A. Polus, J.L. Schofer, A. Ushpiz, Pedestrian flow and level of service, J. Transp. Eng. 109 (1) (1983) 46–56.
[74] D. Helbing, A. Johansson, H.Z. Al-Abideen, Dynamics of crowd disasters: An empirical study, Phys. Rev. E—Statistical, Nonlinear, Soft Matter Phys. 75 (4) (2007) 046109.
[75] J. Wu, S. Ma, Division method for waiting areas on island platforms at metro stations, J. Transp. Eng. 139 (4) (2013) 339–349.
[76] M. Baali, R. Kuipers, R. Coulaud, C. Buisson, C.-W. Palmqvist, Long enough but not too long: a posteriori determination of the dwell time margins from high-resolution passenger flow data, Data Sci. Transp. 7 (2) (2025) 1–19.
[77] L. D'Acierno, M. Botte, A. Placido, C. Caropreso, B. Montella, Methodology for determining dwell times consistent with passenger flows in the case of metro services, Urban Rail Transit 3 (2017) 73–89.
[78] G. Medeossi, A. Nash, Reducing delays on high-density railway lines: London–Shenfield case study, Transp. Res. Rec. 2674 (7) (2020) 193–205.
[79] H.Z. Aashtiani, H. Iravani, Application of dwell time functions in transit assignment model, Transp. Res. Rec. 1817 (1) (2002) 88–92.
[80] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Stat. 22 (1) (1951) 79–86.
[81] S.P. Hoogendoorn, W. Daamen, Pedestrian behavior at bottlenecks, Transp. Sci. 39 (2) (2005) 147–159.
[82] H.E. Nelson, Emergency movement, SFPE Handb. Fire Prot. Eng. (2002).
[83] J. Ye, X. Chen, C. Yang, J. Wu, Walking behavior and pedestrian flow characteristics for different types of walking facilities, Transp. Res. Rec. 2048 (1) (2008) 43–51.
[84] E. Bosina, U. Weidmann, Estimating pedestrian speed using aggregated literature data, Phys. A 468 (2017) 1–29.
[85] I.P.B. Wiggenraad, Alighting and Boarding Times of Passengers at Dutch Railway Stations Doctoral thesis, TU Delft, 2001.
[86] N. Tortainchai, H. Wong, D. Winslett, T. Fujiyama, Train dwell time efficiency evaluation with data envelopment analysis: case study of London underground Victoria line, Transp. Res. Rec. 2676 (3) (2022) 728–739.
[87] F.S. Hänseler, J.P. van den Heuvel, O. Cats, W. Daamen, S.P. Hoogendoorn, A passenger-pedestrian model to assess platform and train usage from automated data, Transp. Res. Part A: Policy Pr. 132 (2020) 948–968.
[88] J. Lee, S. Yoo, H. Kim, Y. Chung, The spatial and temporal variation in passenger service rate and its impact on train dwell time: A time-series clustering approach using dynamic time warping, Int. J. Sustain. Transp. 12 (10) (2018) 725–736.