



**HAL**  
open science

**Long enough but not too long: a posteriori  
determination of the dwell time margins from  
high-resolution passenger flow data**

Mehdi Baali, Ruben Kuipers, Rémi Coulaud, Christine Buisson, Carl-William Palmqvist

► **To cite this version:**

Mehdi Baali, Ruben Kuipers, Rémi Coulaud, Christine Buisson, Carl-William Palmqvist. Long enough but not too long: a posteriori determination of the dwell time margins from high-resolution passenger flow data. *Data Science for Transportation*, In press. hal-04757975

**HAL Id: hal-04757975**

**<https://hal.science/hal-04757975v1>**

Submitted on 29 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Long enough but not too long: a posteriori  
2 determination of the dwell time margins from  
3 high-resolution passenger flow data.

4 Mehdi Baali<sup>1,2\*</sup>, Ruben Kuipers<sup>3†</sup>, Rémi Coulaud<sup>2†</sup>,  
5 Christine Buisson<sup>1†</sup>, Carl-William Palmqvist<sup>3†</sup>

6 <sup>1</sup>LICIT-ECO7, Université Gustave Eiffel/ENTPE, Bron, France.

7 <sup>2</sup>Transilien, SNCF Voyageurs, Saint-Denis, France.

8 <sup>3</sup>Department of Technology and Society, Lund University, Lund, Sweden.

9 \*Corresponding author(s). E-mail(s): [mehdiba5426@gmail.com](mailto:mehdiba5426@gmail.com);  
10 Contributing authors: [ruben.kuipers@tft.lth.se](mailto:ruben.kuipers@tft.lth.se); [remi.coulaud@sncf.fr](mailto:remi.coulaud@sncf.fr);  
11 [christine.buisson@entpe.fr](mailto:christine.buisson@entpe.fr); [carl-william.palmqvist@tft.lth.se](mailto:carl-william.palmqvist@tft.lth.se);

12 †These authors contributed equally to this work.

13 **Abstract**

14 Dwell time is crucial for railway operations corresponding to 20% of the total  
15 travel time in a mass transit context. With this, it is also a source of delays due  
16 to its stochastic nature. One way to ensure the robustness of a timetable is to  
17 add margins and much work has been done on run time margins. On the other  
18 hand, dwell time margins received little attention, except for a few heuristics.  
19 This paper aims to provide a novel method to define dwell time margins. To do  
20 so, we introduce the notion of tight dwell time (dwell without margins) in this  
21 work and compute it from high-resolution passenger flow data. Then, we propose  
22 two novel methods (the cluster method and the quantile method) to estimate it.  
23 Given the access to this data, the method enables the estimation of the tight  
24 dwell time for all the stops while existing heuristics are limited to late trains  
25 and/or few passengers. Our developed method highlights the propensity of exist-  
26 ing heuristics to overestimate what they measure. The estimation, a posteriori,  
27 of dwell time margins thanks to the computed tight dwell time would help design  
28 future timetables.

29 **Keywords:** Railway, dwell time, alighting and boarding time, margins, passenger flow  
30 data.

# 31 1 Introduction

32 The last decade has seen a change in the way the transport sector is viewed, with the  
33 sustainability of the system being called into question (Armstrong & Preston, 2011).  
34 In line with this, efforts are being made to increase the modal share of more sustain-  
35 able modes of transport and induce a shift away from private motorized transport.  
36 Railways are one of the modes that can play a vital role in this process and make  
37 transportation systems more sustainable from an environmental point of view. How-  
38 ever, despite railways providing a green alternative to road transport the potential  
39 environmental benefits will only come to fruition when ridership is sufficiently high  
40 (Givoni et al., 2009), meaning that it is important to make railway an attractive mode  
41 of transport. The attractiveness of railways is affected by the quality of service that is  
42 offered such as the service frequency and on-time performance of railways, along with  
43 travel speed and vehicle tidiness (Mouwen & Rietveld, 2013; van Loon et al., 2011).  
44 These quality aspects of railways present a trade-off, however, since an emphasis on  
45 the frequency of services often comes at the cost of on-time performance (Mouwen &  
46 Rietveld, 2013). To ensure the quality of railways it is important to balance both the  
47 frequency as well as the on-time performance. One way railway operators can directly  
48 influence this balance is through the design of a timetable. The timetable forms the  
49 backbone of railway systems by providing the location and direction of trains at spec-  
50 ified points in time (Goverde, 2005) and indicates the service offered to passengers,  
51 showing both when and how often a train will run.

## 52 1.1 Timetabling

53 Timetable planning, or timetabling, is the task of scheduling when and where trains  
54 will run within a given network, matching the desired train routes to the available  
55 infrastructure (Goverde, 2005). The way a timetable is constructed can directly affect  
56 the level of service for passengers, both in terms of the frequency of the services  
57 offered as well as the ability for trains to run on time. In dense areas, trains operate  
58 with minimal headway, based on the block lengths, characteristics of the train such as  
59 running and braking speeds, and local guidelines (Goverde, 2005; Palmqvist, 2019).  
60 Since conditions such as track adhesion and driving behavior are not always the same,  
61 margins are added to the scheduled times, with the size of these margins depending  
62 on local timetable principles. These margins should be scheduled with care since the  
63 risk of delays remains when scheduling too little time. On the other hand, scheduling  
64 too much time reduces the frequency at which trains can operate. In both cases,  
65 the attractiveness of the railway system suffers. This means that timetabling poses  
66 a balancing act in which it is important to schedule both realistic and appropriate  
67 times for all the different elements that make up the final timetable (Hansen, 2010),  
68 a non-trivial task. Broadly speaking, a timetable for passenger trains can be divided  
69 into two main elements: the run and dwell time of a train. The run time refers to the  
70 time it takes for a train to travel between two subsequent stations, whereas dwell time  
71 refers to the time a train is stationary at a station. The study presented here focuses  
72 on the latter of these, the dwell time.

## 1.2 Importance of dwell times

Dwell times are of interest to study since they make up a large part of the travel time in dense railway networks (about 20% in the Paris region, for example), where trains halt often to allow for passengers to alight and board. Furthermore, dwell times can have a strong effect on the operation of trains in dense areas. Since these trains are bound by the minimum headway the dwell time can become the limiting factor to determine the frequency at which trains can be operated. Running more trains requires shorter dwell times. However, when dwell times are not accurately scheduled this can lead to dwell time delays. Dwell time delays arise when a train is stationary for longer than scheduled (Buchmueller et al., 2008) and although small can accumulate over an entire journey (Christoforou et al., 2020) and cause knock-on delays due to trains occupying a platform for longer than scheduled (Yamamura et al., 2012). A major challenge for railway operators in high-density networks is to define adequate dwell times when designing a timetable: a dwell time which both ensures a feasible timetable (Hansen, 2010) and an optimal use of the available capacity (Goverde, 2005). Issues regarding dwell times will become increasingly important as operators want to increase the frequency of trains operated, often without increasing the available infrastructure. In the Netherlands, for example, there is the ambition to run six trains an hour on some busy corridors (IenW, 2022), putting increased stress on the need for adequate scheduling principles. The time needed for dwell times is hard to predict, however, due to its stochastic nature.

## 1.3 Scheduling of dwell times

Despite the importance of dwell times, current approaches to scheduling dwell times often rely on rules of thumb and the experience from planners (Wiggenraad, 2001; Christoforou et al., 2020). Some studies proposed estimation of dwell times via short-term prediction (Li et al., 2016) but only for short stops with no mandatory departure time. When changes are made to the scheduled dwell time it is generally made a posteriori by increasing dwell time margins when regular dwell time delays appear at certain stations. An example of this approach is described by Matthew Volovski & O’Loughlin (2021) based on the methods proposed in the “Transit Capacity and Quality of Service Manual” used in the US. The dwell time for a given station is estimated based on the historical mean dwell time of the target station or at similar stations. An operational margin of either 15 to 20 seconds or one to two standard deviations of the mean dwell time is then added to this time. This approach will yield increases in dwell time, but it is rare to see a decrease in dwell time when margins are too large. This is partly due to the absence of any method to a posteriori estimate the dwell time margins as most approaches depend on their being a delay, meaning that stations without a delay are effectively ignored. Understanding to what extent margins are in place and their need is important when scheduling dwell times. Knowing the use of margins provides a suggestion of necessary dwell time without margin, referred to as tight dwell time in this work. The tight dwell time is especially relevant for timetabling in heavily utilized railway systems, as the dwell time becomes a limiting factor (see for example the impact of the crowd in simulations done by Luan & Corman (2022)).

116 **Definition.** *Tight dwell time is the necessary time for the dwelling process without*  
117 *any margin.*

118 Other authors have tried to approach this notion, referring to various heuristics  
119 as minimum dwell times which correspond to the minimum time needed to complete  
120 the alighting and boarding process and to depart from a station if everything goes as  
121 fast as possible (Goverde, 2005; Pedersen et al., 2018). Although several dwell time  
122 models that attempt to predict such a minimum dwell time exist, mainly based on  
123 regression models, the potential of such models has not yet been realized (Yang et al.,  
124 2019). For example, testing the well-cited dwell time model by Weston (1989) at a  
125 larger number of stations, Harris (2005) found that the model does not perform well  
126 under high passenger loads.

127 Furthermore, many of the studies focusing on minimum dwell times are limited  
128 by small sample sizes, limited to a small number of stations, or limited to manual  
129 observations (see for example the work done by Puong (2000)). Recent years have  
130 seen an increase in the volume and granularity of the available data (Palmqvist, 2019)  
131 allowing for a more in-depth study into dwell times. At the same time, research into  
132 minimum dwell times has remained scarce, with a recent example being the study  
133 by Cornet et al. (2019) who made use of automatic passenger count data to infer a  
134 minimum dwell time. Moreover, the definition of minimum dwell time indicates that  
135 this is the dwell time corresponding to when everything goes as fast as possible. Such  
136 an assumption is not realistic, that is why we introduce the notion of tight dwell  
137 which instead is an estimate of the necessary dwell time for people to alight and board  
138 properly without any extra time. To the best of our knowledge, no studies have dealt  
139 with a similar notion so far.

#### 140 1.4 Objective of the paper

141 Given the need for an increased understanding of tight dwell times, the objective of  
142 the study presented here is to define a way in which the tight dwell time can be better  
143 estimated for commuter trains. To do so we make use of highly detailed passenger  
144 flow data collected on board suburban commuter trains in the Paris region area and  
145 propose two novel methods to estimate tight dwell times based on high-resolution  
146 passenger flow data. The first method makes use of the time gaps between the flows of  
147 passengers, defining the tight dwell time based on clusters of passengers. The second  
148 method makes use of the cumulative flow of passengers. We then define which of the  
149 proposed approaches is most promising, compare it to methods found in the literature,  
150 and show the usability and precision of our proposed methods. Having insights into  
151 the tight dwell time allows for a posteriori estimate of dwell time margins which would  
152 help the schedule-makers improve future timetables. To summarize, our contributions  
153 are:

- 154 • Introducing the notion of tight dwell time to a posteriori assess the amount of dwell  
155 time margins in a schedule;
- 156 • Proposing two new methods to estimate the tight dwell time;
- 157 • Comparing them to existing methods estimating a close but not identical quantity  
158 (the minimal dwell time).

## 159 1.5 Structure of the paper

160 The remainder of the paper is organized as follows. First, we present a literature  
161 review (Section 2), in which we define the components of dwell time and how dwell  
162 time has been estimated up to now. The formalization of the problem components  
163 is presented in Section 3, and our data is then described (Section 4). Our two new  
164 methods to estimate the tight dwell times are presented in Section 5, along with two  
165 existing methods used to estimate minimal dwell times. A comparison of results from  
166 each method is described in Section 6. This comparison leads us to select the method  
167 based on cumulative flows, and by comparing it further with the two selected methods  
168 from the literature, we show its superiority when sufficiently high-resolution data is  
169 available. An example of the application of the most appropriate method is shown  
170 in Section 7 using data from a branch of *Line N* of the Paris suburban network and  
171 shows that we can even make a door-by-door analysis. A discussion and a conclusion  
172 end the paper.

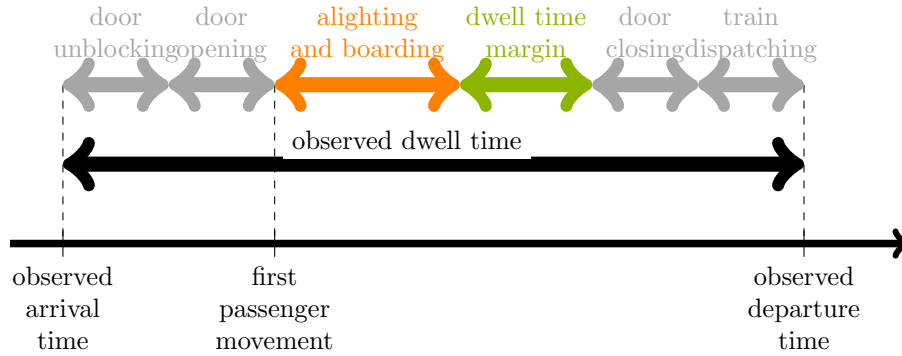
## 173 2 Literature review

174 In this study, we introduce the notion of tight dwell time which is not tackled in the  
175 literature. Yet, the study presented here is supported by several existing works. We  
176 review these works in this section.

### 177 2.1 Components of dwell times

178 Although the dwell time is a single process from a timetabling perspective, it is in  
179 fact made up of different processes. Seriani & Fujiyama (2019) define dwell time as  
180 having both static and dynamic time elements. According to their definition, the static  
181 element of a dwell time consists of mechanical processes such as the door opening  
182 and closing, and the dynamic element consists of the alighting and boarding process.  
183 In line with this, Buchmueller et al. (2008) state two main processes make up the  
184 dwell time, these being the passenger service time and the train dispatching time.  
185 The authors further divide dwell time into five different sub-processes, including the  
186 door-unblocking, door opening, boarding/alighting, door closing, and train dispatching  
187 process. Figure 1 illustrates this division of the dwell time.

188 The time needed to complete the static elements of the dwelling process is depend-  
189 ent on the design of the train. Trains can be equipped with sliding extensions to  
190 allow for a level entry, for example. The time needed to open the door is longer when  
191 such an extension is in place (Buchmueller et al., 2008). Although considered to be a  
192 static element of dwell times, the time needed to open a door can be dynamic when  
193 the doors do not open automatically and passengers are required to request a door  
194 opening (Harris, 2015). The time needed to complete the dynamic element of dwell  
195 times is governed by the alighting and boarding time (Goverde, 2005). Since trains  
196 have multiple doors there are multiple alighting and boarding processes at the same  
197 time, and it is the door where this takes the longest which defines the alighting and  
198 boarding time making it the "critical door" (Buchmueller et al., 2008; Coulaud et  
199 al., 2023). Note that the critical door is specific to the station, the time of the day,



**Figure 1** Dwell time components as presented by Buchmueller et al. (2008)

200 the rolling stock, and so on. This adds another source of variability, which concurs in  
 201 justifying the necessity for margins.

## 202 2.2 Margins and robustness

203 When designing a timetable it is important to include realistic running and dwell  
 204 times to ensure that the final timetable is feasible (Hansen, 2010). One way to ensure  
 205 the feasibility of the scheduled times is to include some form of margins in both the  
 206 running and dwell times. Here, margins refer to time supplements added on top of  
 207 the scheduled times to allow for any fluctuations, be it a result of driver behavior or  
 208 delays (Andersson et al., 2011; Goverde, 2005). These margins are important since it is  
 209 unlikely that the actual operation will follow the scheduled times. Adding margins to  
 210 running times is common in timetables (Palmqvist, 2019), and are usually a percentage  
 211 of the nominal running time. This percentage is based on local planning principles,  
 212 with seven percent being the norm in the Netherlands (Goverde, 2005), three percent  
 213 in Sweden (Palmqvist, 2019), and five percent for the suburban trains operated in  
 214 Paris. Other margins can be added at important locations, also known as nodes, and  
 215 is common practice in Sweden (Palmqvist, 2019) and Switzerland (Vromans, 2005),  
 216 for example, and is also common practice in France.

217 Having these margins in place helps with the robustness of a timetable, meaning  
 218 that small deviations from the scheduled time do not lead to a delay (Dewilde et  
 219 al., 2011). Adding margins is a balancing act, however, where too many, or too large  
 220 margins will result in unnecessarily long travel times. On the other hand, too few, or  
 221 too small margins can lead to robustness issues. To ensure a robust timetable, planners  
 222 need thus to measure the robustness of a timetable and modify the timetable in such  
 223 a way that robustness is maintained whilst keeping other planning objectives in mind  
 224 (Andersson et al., 2013).

## 225 2.3 Estimating dwell time margins

226 As mentioned, margins can be added to both the running and dwell time of trains.  
 227 Running time margins are relatively straightforward to calculate since the nominal  
 228 running time can be calculated based on a feasible speed-distance profile over an open

229 stretch of track for a train, taking the track alignment characteristics into account  
230 (Goverde, 2005). The nominal running time can then be compared to the realized  
231 operation of a given set of trains to measure the margins present in the timetable. By  
232 contrast, dwell time margins are less straightforward to calculate given the stochastic  
233 nature of dwell times. The alighting and boarding time, especially, makes the actual  
234 dwell time subject to high variability (Buchmueller et al., 2008), and comparing the  
235 scheduled and realized dwell times is not trivial.

236 Only a limited number of methods have been proposed in the past to measure the  
237 margin present in dwell times. One way to do so is to make use of situations where a  
238 train arrives with such a delay that it enters the station after its scheduled departure  
239 time. An example of this approach was explained by Pedersen et al. (2018). Coulaud  
240 et al. (2023) also highlighted that late train dwell times are impacted at the first order  
241 by the passenger flows which suggests a reduction of margin for late trains. In such  
242 cases, trains are supposed to depart as soon as possible, and the dwell time is likely to  
243 only include the alighting and boarding time. Yet, none of these approaches quantified  
244 the margins for late trains.

245 Moreover, some authors (Cornet et al., 2019) criticized this kind of approach high-  
246 lighting that some late trains could not leave the station right after the alighting and  
247 boarding process was completed, due to signalization and dispatching decisions for  
248 instance. Daamen et al. (2009) and Kecman & Goverde (2013) proposed methods to  
249 identify route conflicts that could be used for filtering out trains without conflict. For  
250 example, Kecman & Goverde (2015) presented a machine-learning approach to pre-  
251 dict dwell times using this filtering method. Apart from route conflicts, the driver's  
252 behavior and his sensitivity to how much time to wait before leaving may impact such  
253 estimation of the alighting and boarding time. Another approach to overcome these  
254 problems is proposed by Cornet et al. (2019) taking advantage of passenger flow data  
255 instead. Having access to detailed passenger count data (i. e. the total number of peo-  
256 ple alighting and boarding the train and the load of the train at every stop) Cornet et  
257 al. (2019) defined a measure they call the reduced passenger flow. A minimum dwell  
258 time is subsequently obtained as a function of the reduced passenger flow. Whilst  
259 promising, the two approaches described above are only defined for late trains and do  
260 not allow to capture the variability in the flow of both alighting and boarding passen-  
261 gers. Moreover, only the study of Coulaud et al. (2023) used door-by-door counting  
262 data but focused on dwell time prediction. In this work, we are presenting a way to  
263 estimate tight dwell times for any stop, a metric that captures the variability of the  
264 passenger flow. From there, one can then deduce the dwell time margin by comparing  
265 the tight dwell time to the observed dwell time.

### 266 **3 Formalization of the problem components**

267 In this section, we are first introducing a general mathematical framework for counting  
268 data and dwell times components data.



Variable	Index
Train number	$k$
Station	$s$
Date	$d$
Train door index	$j$
Time (relative to the opening of the door)	$t$

**Table 1** Indexes.

Quantity	Expression
Total volume of passenger for one line	$\sum_{k,s,d,j,t} N_{k,s,d,j,t}$
Passengers' volume in one day/one line	$\sum_{k,s,j,t} N_{k,s,d,j,t}$
Annual station volume	$\sum_{k,d,j,t} N_{k,s,d,j,t}$
Annual volume for a train and a station	$\sum_{d,j,t} N_{k,s,d,j,t}$
Daily station volume	$\sum_{k,j,t} N_{k,s,d,j,t}$
Volume by stop	$\sum_{j,t} N_{k,s,d,j,t}$
Volume by door	$\sum_t N_{k,s,d,j,t}$
Alighting or boarding at time $t$ (relative to the latest opening of the door)	$N_{k,s,d,j,t}$

**Table 2** Examples of levels of resolution and their formalized expression.

### 3.1 Hierarchy of resolutions in passengers flow data

In this section, we introduce the different levels of resolution (and the related indexes) that may exist concerning the available passenger flow data. The highest resolution for counting passengers that either alight or board a train is the movement of passengers through a door. A passenger is defined as someone who alights or boards a train once. In practice, this means that a passenger who has boarded a train and then alights the train later will be counted twice, once as a boarding and once as an alighting passenger. Count data on this resolution does, however, not provide information on when and where the event took place. To be able to do so, four pieces of information are needed: the train number  $k$ , the station  $s$ , the day  $d$ , and the door  $j$ . An additional fifth piece of information is required to be able to study dwell time components, this being the time when a passenger crosses the door relative to the opening of the door  $t$ . Table 1 sums up these notations.

The highest resolution of data can be noted as  $N_{k,s,d,j,t}$ , and provides information on the number of passenger movements for train  $k$ , at station  $s$ , on date  $d$ , at door  $j$ , and time  $t$  after the opening of the door. Lowering this resolution is often synonymous with aggregating  $N_{k,s,d,j,t}$  over several indexes that correspond to operational quantities. Table 2 shows several examples of levels of resolution. For example, summing  $N_{k,s,d,j,t}$  over time ( $t$ ) and door ( $j$ ) results in the volume of passengers per stop; and summing  $N_{k,s,d,j,t}$  over time ( $t$ ), door ( $j$ ), and train number ( $k$ ) results in the daily station volume. Summing over all the indexes means getting the total volume of passengers for the studied line in the studied period.

Although higher resolutions enable a more precise analysis, it is difficult to access such data in practice. Indeed, three kinds of restrictions may exist to access higher resolutions of passenger flow data:

- 294 1. **Data collection.** If the data is collected manually, it is not possible to be precise  
295 enough to get the time  $t$  of the passengers' movement relative to the opening of the  
296 door;
- 297 2. **Data storage.** If the data is collected automatically, the data has to be transmitted  
298 from the train to a server and stored. The highest resolution makes the data heavier  
299 and the storage on-board and transmission off-board are either complicated or  
300 expensive which is dissuasive for many transport companies;
- 301 3. **Competition between companies.** Knowing precisely the passengers' move-  
302 ments is a critical piece of data for railway operation and, provided they have higher  
303 resolution data available, some companies may consider it confidential.

304 As a consequence, most of the work done using counting data so far has used  
305 lower-resolution data. Figure 2 illustrates the data hierarchy by showing the levels  
306 of resolution and associated use cases and works. At a macroscopic scale, the Global  
307 Transport Investigation (DRIEA, 2022) shows, among other information, the evolution  
308 of the total volume of passengers in the Paris region every decade. The annual station  
309 volume is also interesting to model, especially for new stations or new lines, the four-  
310 step model is the most common tool to do so (McNally, 2007). The passenger volume  
311 of one day (or one time period) can be a useful piece of data to predict disruption  
312 for example (Yap & Cats, 2021). When only manual counting data is available, the  
313 annual volume for a train and a station is the most appropriate piece of data to build  
314 timetables (Brethomé, 2018). Besides, the daily station prediction (De Nailly, 2023)  
315 does not need more than daily station volumes. The volumes by stop helped improve  
316 the robustness of timetables (Cornet et al., 2019; Medeossi & Nash, 2020; Christoforou  
317 et al., 2020; Kuipers & Palmqvist, 2022) while the volume by door is a precious piece  
318 of data to model dwell times (Coulaud et al., 2023; Kuipers, 2024). The highest level of  
319 resolution has not been tackled in the literature, to the best of our knowledge. Yet, it  
320 could offer various implications to better understand the dynamic of the alighting and  
321 boarding process. Among them, this work only focuses on tight dwell time estimation.

### 322 3.2 Dwell time components

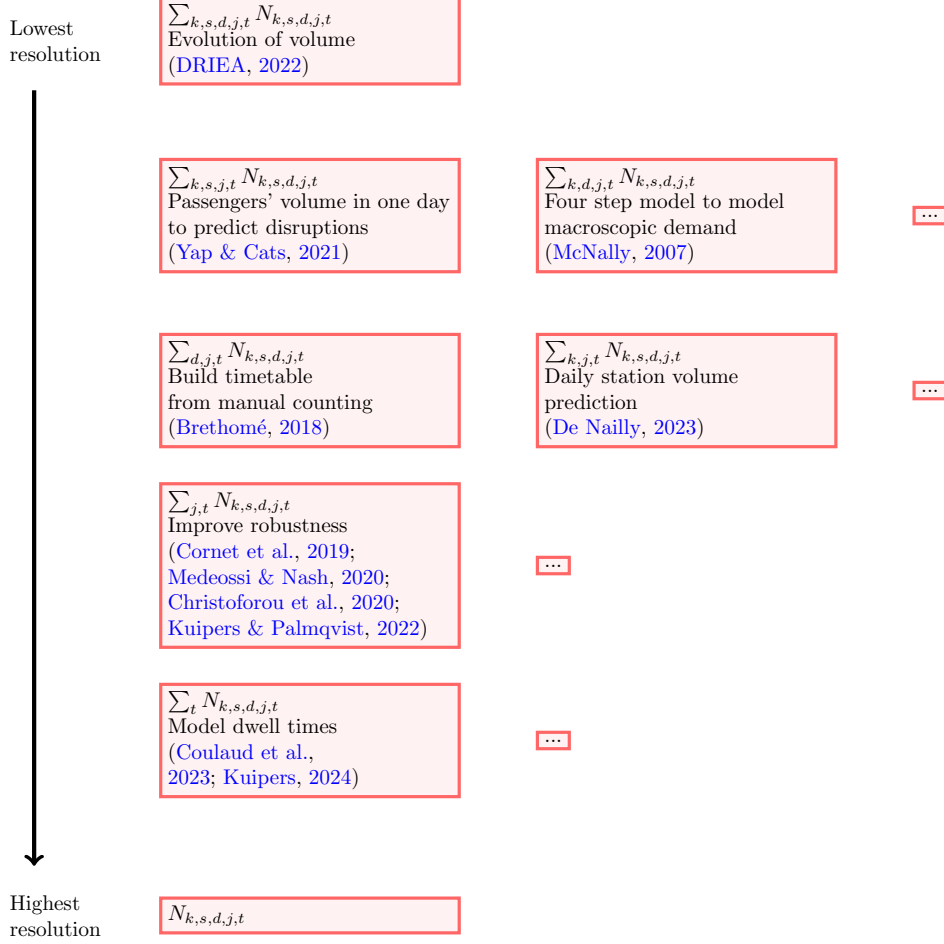
323 The observed dwell time ( $DT$ ) of a train is composed of several components: the  
324 technical time ( $TT$ ), the alighting and boarding time ( $ABT$ ), and some dwell time  
325 margin ( $DTM$ ):

$$326 \quad DT = ABT + DTM + TT.$$

327 The resolution of the data we have access to enables the design of new methods to  
328 estimate directly the alighting and boarding time. In Section 5, two new methods that  
329 we have developed are presented. We can then define the tight dwell time ( $TDT$ ) as  
330 the dwell time without margin as follows:

$$331 \quad TDT = ABT + TT$$

332 We can suppose the technical time ( $TT$ ) to be constant among the doors and the  
333 stops (Buchmueller et al. (2008) showed some variability but was less important than  
334 the one observed for the other variables). By contrast, the alighting and boarding



**Figure 2** Use cases of passenger flow data by level of resolution, where  $N$  is the number of passengers,  $k$  the train number,  $s$  the station,  $d$  the day,  $j$  the door, and  $t$  the time when the passenger cross the door (relatively to the opening of the door). The omission marks correspond to other ways to aggregate the passenger flow. Existing works can be found for each level of resolution except the highest.

335 time can vary among the stops and the doors for a given stop and the margin can  
 336 differ between stops as well. The alighting and boarding time is the necessary time for  
 337 all the passengers to either alight or board the train. As an example, the time for a  
 338 passenger arriving late at the station and boarding right before the departure of the  
 339 train is included in the alighting and boarding time while a long period between two  
 340 passengers' movements is not included. We define the alighting and boarding time for  
 341 a given door  $j$  as a door alighting and boarding time ( $DABT_j$ ), and the margin as a

342 door dwell time margin ( $DDTM_j$ ). We then have:

$$343 \quad ABT = \max_j DABT_j$$

$$344 \quad DTM = \min_j DDTM_j$$

### 346 3.3 Cornet's minimum dwell time definition

347 [Cornet et al. \(2019\)](#) introduced the notion of minimum dwell time ( $MDT$ ) for a  
 348 reduced passenger flow ( $p$ ), for a fixed station  $s_0$ .  $p$  is obtained by doing a PCA on flow  
 349 data (number of alighting  $A$ , number of boarding  $B$ , load  $L$ ), and is then normalized  
 350 to have a value between 0 and 1.

$$351 \quad p = \frac{\beta_A A + \beta_B B + \beta_L L}{\max \beta_A A + \beta_B B + \beta_L L}$$

352 where  $\beta_X$  represents the weights obtained for the first principal component of the PCA.  
 353 Intuitively,  $p$  represents the size of the passenger flow such that  $p = 0$  corresponds to  
 354 a train with almost no passenger exchange and  $p = 1$  to a train with a busy and messy  
 355 alighting and boarding process. The minimum dwell time for a fixed reduced flow  $p_0$   
 356 is obtained as follows:

$$357 \quad MDT(p_0, s_0) = \min_{k,d \text{ s.t. } p=p_0} DT_{k,s_0,d}$$

358 In the remainder of this paper, we estimate  $MDT$  and  $TDT$  with different methods,  
 359 the notation  $\hat{X}$  is used to estimate any variable  $X$ . All the notations are summarized  
 360 in Table 3. We will then compare the results of  $\widehat{TDT}$  computed with our new methods  
 361 with  $\widehat{MDT}$  as estimated with Cornet's method. Yet, we do not estimate the same  
 362 variables. However, by adding the hypotheses that  $\widehat{ABT}$  and  $\widehat{DTM}$  are independent  
 363 and that  $\lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{DTM}_{k,s_0,d} \right) = 0$ , we can demonstrate that:

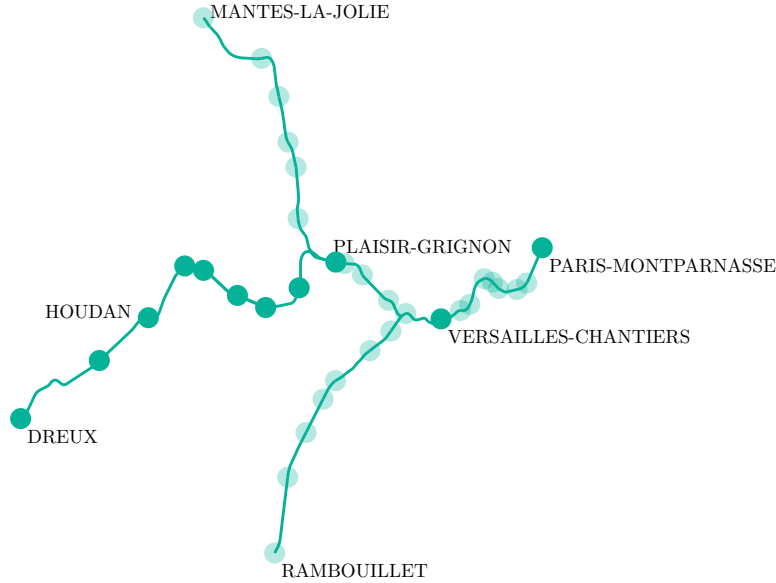
$$364 \quad \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{TDT}_{k,s_0,d} \right) = \lim_{|\{k,d\}| \rightarrow \infty} \widehat{MDT}(p_0, s_0).$$

365  $ABT$  is a flow variable determined by how passengers alight and board the train  
 366 and how many they are, whereas  $DTM$  is an operational variable and is determined  
 367 by the departure time of a train. Thus, the independence between the estimation of  
 368 these variables is a reasonable hypothesis. As far as the limit of the minimal margin  
 369 is concerned, a case may exist where the train departs as soon as the alighting and  
 370 boarding process is finished (assuming an infinite dataset) which corresponds to  $DTM$   
 371  $= 0$ .

372 This proposition (demonstrated in Appendix) makes it possible to perform a com-  
 373 parison between  $\widehat{MDT}$  as defined by [Cornet et al. \(2019\)](#) and  $\widehat{TDT}$ , provided enough

Variable	Notation
Dwell Time	$DT$
Alighting and Boarding Time	$ABT$
Dwell time margin	$DTM$
Technical Time	$TT$
Local Alighting and Boarding Time (door level)	$DABT$
Local Dwell Time Margin (door level)	$DDTM$
Minimum Dwell Time (for fixed $p$ and $s$ )	$MDT(p, s)$
Tight Dwell Time	$TDT$

**Table 3** Summary of the notations. All these variables are times with seconds as units. We introduce the tight dwell time which is the dwell time without margin.



**Figure 3** Map of *Line N*. All dots are stations served by *Line N*. Opaque dots correspond to the services of branch Paris-Dreux that we will tackle in the sequel.

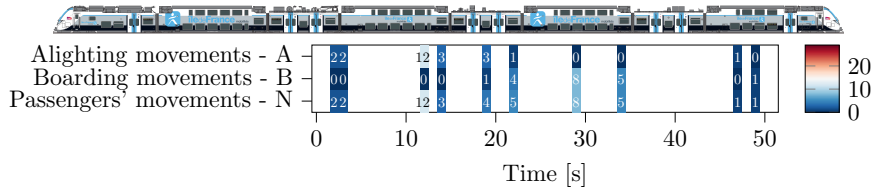
374 amount of data, that we estimate based on the methods presented in the sequel of this  
375 study.

## 376 4 Presentation of our data and perimeter

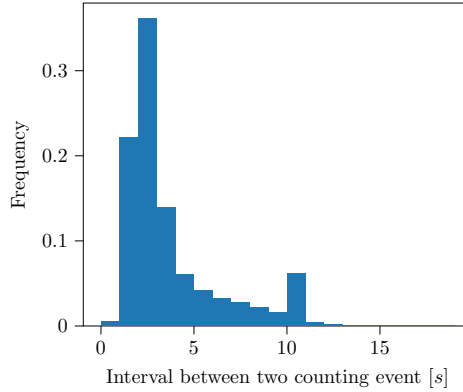
377 In this section, we are presenting the perimeter and the specific data that we used in  
378 this study.

### 379 4.1 High resolution passenger flow data used in this study

380 The data used in this study originates from *Line N* of the Paris suburban trains net-  
381 work, shown in Figure 3 and has been collected over six months from September 2022  
382 to March 2023. *Line N* consists of three branches and various services are operated on



**Figure 4** Regio2N design for *Line N* of the Paris suburban trains network (top). Data at the counting event scale is obtained via the infrared sensors of the Regio2N (bottom). Counting events are the number of passengers' movements since the last counting event. Two counting events are separated randomly by several seconds.



**Figure 5** Distribution of interval between two counting events. The average is 3s with a standard deviation of 3s. A non-explained peak is observed at 10s.

383 each branch (omnibus and semi-directs). The rolling stock in use is the Regio2N, il-  
 384 lustrated in Figure 4. Each carriage has eight doors which are 1.60m wide and require  
 385 passengers to press a button to initiate the opening procedure. The technical time is  
 386 supposed deterministic and uniform at 7.5s (as reported in the technical documenta-  
 387 tion). The services on this line are operated with both single units and double units.  
 388 *Line N* is an intermediary line in terms of passenger volumes (around 125,000 people  
 389 on weekdays) but it is the most used among lines equipped with Regio2N rolling stock.

390 The rolling stock in operation is equipped with an automatic passenger count sys-  
 391 tem at each door which counts both the number of alighting and boarding passengers  
 392 using infrared sensors. Every several seconds, a signal is sent to an onboard computer  
 393 that registers the cumulative number of passenger movements. In addition to this, the  
 394 system registers whether the train doors are open or closed. It also provides informa-  
 395 tion on the train number and the name of the stop. Several stops ( $k$ ,  $s$ ,  $d$ ) (around  
 396 30%, randomly distributed) are not counted by the sensors which leads to gaps in the  
 397 data. We are, therefore, limited to focusing on the stops where the data is available.

398 The study presented here (see Section 6) focuses on *Versailles-Chantiers*, which  
 399 is the main station in *Line N* (except for Paris terminus) both in terms of passenger  
 400 traffic (nearly 15 000 passengers boarding per day) and in terms of train traffic (more  
 401 than 200 trains a day on weekdays). To compare the methods, we select only trains

402 toward Paris. A more general analysis is performed in Section 7 where we consider all  
403 the services on *Dreux* branch. Two trains per hour during peak hours and one train  
404 per hour during off-peak hours are operated on this branch. This branch undergoes  
405 two convergences with the other branches of the line: the first between *Villiers-*  
406 *Neauphle-Pontchartrain* and *Plaisir-Grignon*, the second between *Plaisir-Grignon* and  
407 *Versailles-Chantiers*.

408 The automatic passenger count data makes it possible to have access to data close  
409 to the highest resolution we presented previously. The data is disaggregated in terms  
410 of  $k, s, d, j$  and partially aggregated on  $t$  (as shown in Table 1). Indeed, for each  
411 quadruplet  $(k, s, d, j)$ , we have a list of counting events. A counting event means that at  
412 a given time  $t'$ ,  $N_{k,s,d,j,t'}$  passengers have alighted/boarded the train since the latest  
413 counting event (or the opening of the door in the case of the first counting event).  
414 Figure 4 shows an example of our data.

415 The time separating two counting events is truly random. Figure 5 shows the  
416 distribution of the interval between two counting events. The average is  $3s$  and the  
417 standard deviation is  $3s$ . Most of the values are below  $10s$  but one may notice the  
418 important peak at  $10s$ . We have no explanation so far for this peak but this will  
419 have some implications on the tight dwell time estimation. The length of the interval  
420 between two counting events provides an estimate of accuracy on time: the indicated  
421 number of passengers have boarded or alighted between this counting event and the  
422 previous one. Then the precise time of one passenger's movement is at worst included  
423 in the interval, which provides quite a pessimistic uncertainty. As an uncertainty for  
424 time, we will consider the root quadratic mean (i.e. a counting event time is at worst  
425  $4.5s$  late on average).

426 In terms of number of passengers, the data quality assessment is more challeng-  
427 ing. The daily alighting and boarding volumes for a given train number are compared  
428 to determine the accuracy of the system. For a given train  $k$  at a given date  $d$ , the  
429 total number at all stops of people alighting shall be the same as the total number  
430 of people boarding. In the case of the presented data, the total number of alighting  
431 differs by 5% on average (3% in median) from the total number of boarding which  
432 gives an insight into a decent data quality. In addition to this, a survey has been per-  
433 formed to compare the infra-red counting data aggregated at the stop scale to manual  
434 counting. Assuming manual counting to be the truth, the survey highlighted a bias  
435 of  $+0.13$  people for boarding and  $+0.24$  for alighting. Those biases, though small,  
436 might be explained by double counting of people crossing the door slowly or jerkily.  
437 A standard deviation of 4.4 people is observed for boarding and 3.5 people for alight-  
438 ing, probably partly corresponding to a confusion of the sensor between alighting and  
439 boarding. In the remainder of this work, we will consider any passenger's movement  
440 (amalgamating alighting and boarding), this will partly mitigate sensors' confusion.  
441 Indeed, concerning passenger's movements, the standard deviation compared to man-  
442 ual counting is 4.4 people. The error is still significant but it is not a deterrent for  
443 practical use (the threshold imposed by the public transport authority being 5 people  
444 in standard deviation in the Paris region). Moreover, the contribution of this work is  
445 above all methodological and the data quality is enough to deal with methods.

## 446 4.2 Dwell time measurement

447 For dwell time measurement, several kinds of sensors measure the observed dwell times  
448 via automatic vehicle localization (AVL). In this study, we used data from the infrared  
449 sensors. These sensors indicate the time when the door is fully open and the time  
450 when the door is fully closed. Using the door unlock and locking times, as opposed to  
451 the time between wheel stop and start ensures that the dwell time we use is related  
452 to the movement of passengers and excludes dispatching time.

## 453 5 Presentation of methods to estimating tight dwell 454 times

455 As stated before, part of the objective of the study presented here is to compare our  
456 two methods to estimate the tight dwell time to heuristics proposed in the literature.  
457 These methods are presented in this section, starting with a description of the two  
458 heuristics proposed in the literature.

### 459 5.1 A simple heuristic: observed dwell times for late trains

460 One way to determine the tight dwell time is to make use of situations where the  
461 train enters the station with such a delay that it arrives after its scheduled departure  
462 time. An example of this approach was explained by [Pedersen et al. \(2018\)](#). They  
463 hypothesized that a late train would only dwell for the shortest time needed, i.e. the  
464 time needed for the alighting and boarding process to be completed. This is commonly  
465 done to recover some of the delay time. The dwell time margin is then at least partly  
466 consumed as shown by [Coulaud et al. \(2023\)](#). In the study presented here, we will  
467 compare the dwell time in this situation to the tight dwell time that we will estimate  
468 with our new method.

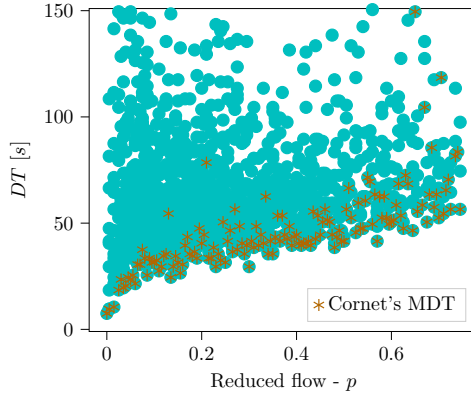
469 One particularity of this method is that it does not need any passenger flow data  
470 which makes it an attractive method: it only relies on the scheduled and actual arrival  
471 and departure times. Nevertheless, it does come with a drawback; [Cornet et al. \(2019\)](#),  
472 for example, criticized this approach highlighting that some delayed trains could not  
473 leave the station right after the alighting and boarding process was complete due to  
474 signalization and dispatching decisions for instance. This could lead to situations where  
475 dwell time is significantly greater than the tight dwell time. Some works proposed  
476 corrections to improve this approach, filtering route conflict ([Daamen et al., 2009](#);  
477 [Kecman & Goverde, 2013, 2015](#)) but for the sake of simplicity, we will not further  
478 discuss these route conflict corrections in this paper.

### 479 5.2 A heuristic based on passenger volumes: Cornet's 480 minimum dwell time

481 Given the influence of passengers on dwell times, [Cornet et al. \(2019\)](#) proposed to use  
482 passenger flow data to estimate a minimum dwell time. The workflow proposed is the  
483 following:

- 484 1. Select one station and one direction;





**Figure 6** Illustration of how to get Cornet’s minimum dwell time ( $\widehat{MDT}$ ) from the observed dwell time *Versailles-Chantiers* toward Paris.  $\widehat{MDT}(p)$  being the minimum observed dwell time for  $p$ ,  $p$  represents the size of the passenger flow.

- 485 2. Perform a principal component analysis (PCA) on the number of alighting at one
- 486 stop, the number of boarding at one stop, and the load of the train to define  $p$  (the
- 487 normalized first principal component) called reduced passenger flow;
- 488 3. Consider only trains for which the observed departure time is strictly greater than
- 489 the scheduled departure time;
- 490 4. Separate all the stops by their values  $p$  in 200 windows ( $[0,0.005];[0.005,0.01] \dots$ );
- 491 5. Get the minimum dwell time  $\widehat{MDT}$  as the minimal observed dwell time among the
- 492 window of  $p$  in which the stop is.

493 In this case,  $p$  is an illustration of the relative importance of the passenger flow

494 and is a value between 0 and 1 with  $p = 0$  meaning that there are no passengers,  $p =$

495 1 indicating the maximum flow encountered in the dataset.

496 In line with the available resolution of data ( $\sum_{j,t} N_{k,s,d,j,t}$ ), Cornet et al. (2019)

497 assumed that the flow was uniformly distributed among the doors, something which

498 is not likely. Given that the data available for our study provides information on the

499 distribution of passengers on a door-by-door level it is possible to include this aspect.

500 Another adaptation to the method proposed by Cornet et al. (2019) is the exclusion of

501 the load factor of the train. The line studied here is not heavily used and it is possible

502 to assume that standing passengers are rarely present given the seating capacity of

503 the trains.

504 These adjustments mean that we changed step 2 and step 4 of the workflow presented

505 above. In step 2, we performed a PCA on the number of people alighting at one

506 door and the number of people boarding at the same door  $j$  ( $p_j = \frac{\beta_A A_j + \beta_B B_j}{\max \beta_A A_j + \beta_B B_j}$ ). In

507 step 4, we consider for each stop the maximum value  $p$  among the doors before separat-

508 ing the stops. Figure 6 shows the last step of this framework for *Versailles-Chantiers*

509 toward Paris.

510 This method is based on the hypothesis that  $MDT$  entirely depends on the reduced

511 passenger flow  $p$ . However, one may conjecture the existence of instabilities in the

512 alighting and boarding time regarding the passenger flow such as slower passengers  
513 and passenger boarding before the end of the alighting process.

### 514 **5.3 Tight dwell time estimation using high-resolution** 515 **passenger flow data**

516 The following sections present two methods to estimate the tight dwell time, based on  
517 high-resolution passenger flow data (presented in Section 4.1). Both methods estimate  
518 the door alighting and boarding time ( $\widehat{DABT}$ ) based on the profile of the passenger  
519 flow. The first method is referred to here as the "cluster method" in which the alighting  
520 and boarding time at a door is determined as a function of clusters of passengers. The  
521 second method is referred to here as the "quantile method" in which the necessary time  
522 for a specified quantile of the passenger flow to perform the alighting and boarding  
523 process is considered.

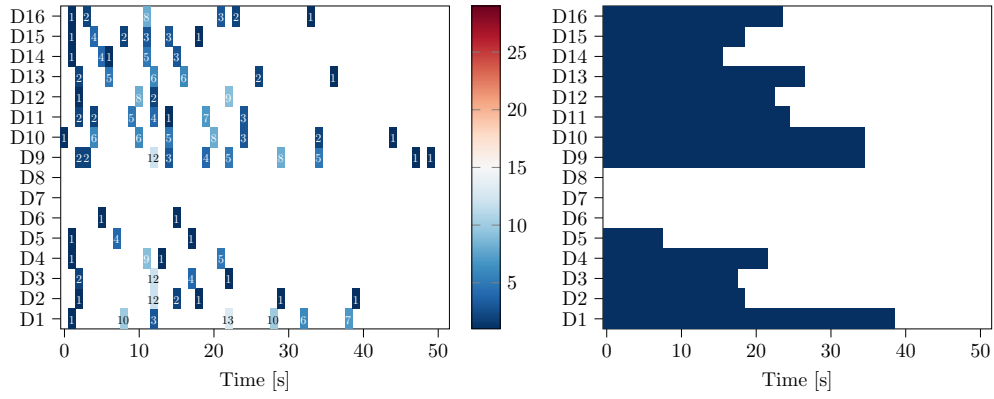
#### 524 **5.3.1 Cluster method**

525 The cluster method is inspired by an existing method for the estimation of the alight-  
526 ing and boarding time at the door level (Wiggenraad, 2001). The main challenge to  
527 estimate  $\widehat{DABT}$  apart from dwell time margins is to distinguish the main flow of  
528 passengers from passengers boarding just before the doors close. Such late arriving  
529 passengers, or people remaining on the platform waiting for the alarm to board, may  
530 bias our understanding of the alighting and boarding process. Then, one needs to make  
531 the difference between isolated movements and cluster movements. For Wiggenraad  
532 (2001), a cluster is the movement of a group of passengers separated by strictly less  
533 than four seconds (the cluster size parameter,  $\Delta_t$ ). By contrast, an isolated movement  
534 would be the movement of one person separated by more than four seconds from the  
535 other passengers.

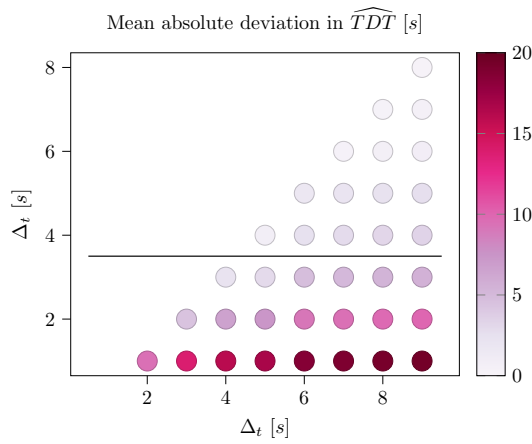
536 Such a protocol could be implemented without modification if the data had the  
537 highest resolution. In our case, we deal with counting events, so we need to adjust the  
538 existing method. Thus, we take the following definition: two counting events are in the  
539 same cluster if there are less than  $\Delta_t$  seconds by passengers in the later counting event  
540 relative to the time between the two counting events. We define the estimated  $\widehat{DABT}$   
541 as the duration of the first cluster. Figure 7 illustrates the process of estimating  $\widehat{DABT}$   
542 with the cluster method. As explained in Section 3, to get  $\widehat{ABT}$  one needs to take the  
543 maximum  $\widehat{DABT}$  among the doors of the train for this stop ( $\widehat{ABT} = \max_j \widehat{DABT}_j$ ).

544 Therefore, the estimated tight dwell time is  $\widehat{TDT} = \widehat{ABT} + TT$ .

545 The value of the cluster size parameter ( $\Delta_t$ ) needs to be chosen carefully. Indeed,  
546 a too-low value for  $\Delta_t$  would lead to splitting clusters in isolated movement, resulting  
547 in an underestimation of  $\widehat{TDT}$ . A too-high value for  $\Delta_t$  would imply considering too  
548 many passengers in the cluster and part of the dwell time margin would be included in  
549 the estimated  $\widehat{TDT}$ . As an illustration, Figure 8 shows the mean absolute deviation in  
550  $\widehat{TDT}$  for various  $\Delta_t$  values (among all the trains toward Paris at *Versailles-Chantiers*).  
551  $\Delta_t=4s$  chosen by Wiggenraad (2001) is a decent value since all values below  $\Delta_t \leq 4s$



**Figure 7** Heatmap of counting events and the number of passengers' movements in each counting event for each door (left), induced clusters representing the main flow of passengers (right), time 0 is the opening of the doors.

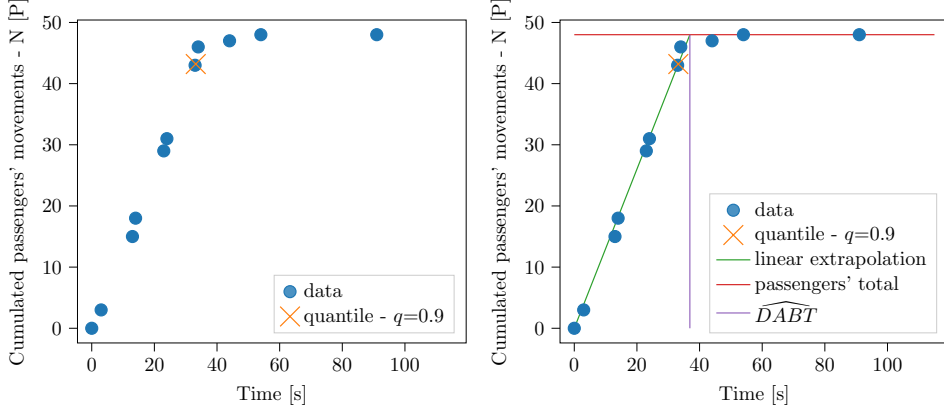


**Figure 8** Heatmap of the mean absolute deviation in  $\widehat{TDT}$  computed with the cluster method varying  $\Delta_t$  two by two.  $\Delta_t \geq 4s$  corresponds to smaller deviations. Choosing  $\Delta_t$  in this range leads to robust results.

552 correspond to significantly more important deviation (meaning important sensitivity  
 553 to the parameter that we want to avoid here). From this analysis, the estimations are  
 554 rather close for any  $\Delta_t \geq 4s$ , with less than  $3s$  of mean deviation (compared to tight  
 555 dwell times around  $35s$  on average). To confirm this statement, we are focusing on  
 556  $\Delta_t=4s$  and  $\Delta_t=6s$ , in the sequel.

### 557 5.3.2 Quantile method

558 Another way to deal with isolated movements is to consider the necessary time for a  
 559 quantile  $q$  of the passenger flow to perform the process. This necessary time then needs  
 560 to be expanded to take into account all the passengers proportionally to the portion  $q$



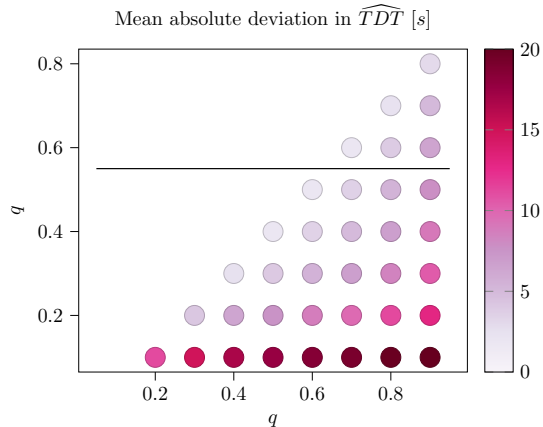
**Figure 9** Cumulative passengers (alighting and boarding) against time (blue dots) and quantile  $q=0.9$  (the orange cross), (left)  $\widehat{DABT}$  computation by adding a time proportional to the number of remaining passengers to the time of quantile  $q=0.9$ , the linear extrapolation is the green line, the passengers' total is the red line and  $\widehat{DABT}$  corresponding to the intersection of both lines is in purple (right).

561 of them. Figure 9 illustrates this method for  $q=0.9$ . On the left of the figure, we can see  
 562 that the last passengers' movements are realized much later than the other passengers'  
 563 movements. Then, we extrapolate the necessary time for this quantile  $q$  (orange cross)  
 564 to obtain the time per passenger to alight/board the train (i.e. the slope of the green  
 565 line). We then take  $\widehat{DABT}$  (purple line) as the product of this time per passenger  
 566 by the total number of passengers which corresponds to the intersection between the  
 567 green line and the red line. The idea behind this method is to give every passenger  
 568 a sufficient amount of time to alight/board the train without adding too much time  
 569 for the slowest passengers. Once again, this method defines an estimation of  $\widehat{DABT}$ ,  
 570 and one needs to take the maximum among the doors and add the technical time to  
 571 obtain the desired estimation of  $\widehat{TDT}$ .

572 The choice of  $q$  is important. As  $q$  is an extrapolation parameter, choosing  $q \leq 0.5$   
 573 would mean more than half of the flow is extrapolated. Similarly, a value of  $q$  close to  
 574 1 means nearly no extrapolation (considering people who are not in the main flow).  
 575 Figure 10 shows the mean absolute deviation in  $\widehat{TDT}$  for various  $q$  values (among all  
 576 the trains toward Paris at *Versailles-Chantiers*). Small mean absolute deviations (less  
 577 than 5s) are observed for  $q$  between 0.6 and 0.9 to be compared with tight dwell times  
 578 of the order of 35s on average. It means robust results in this range. In the remainder  
 579 of this paper, we will focus on  $q=0.6$  and  $q=0.8$  and compare this method to the three  
 580 already presented.

## 581 6 Results: comparison of the methods

582 In this section, we show the comparisons between the four methods presented previ-  
 583 ously. In the following subsections, the most restrictive sample is chosen to compare



**Figure 10** Heatmap of the mean absolute deviation in  $\widehat{TDT}$  computed with the quantile method varying  $q$  two by two. The method is relevant for  $q$  over 0.5 and less than 1. Little deviation is observed in this range. Choosing  $q$  in this range should lead to robust results.

Method	Subset of the data used	Sample size	Subsection
Quantile method	All	2381	6.1
Cluster method	All	2381	6.1
Cornet's minimum dwell time	Trains leaving later than their theoretical departure time windows having several values of $p$	1392	6.3
Observed dwell time for late trains	Trains arriving later than their theoretical departure time	467	6.2

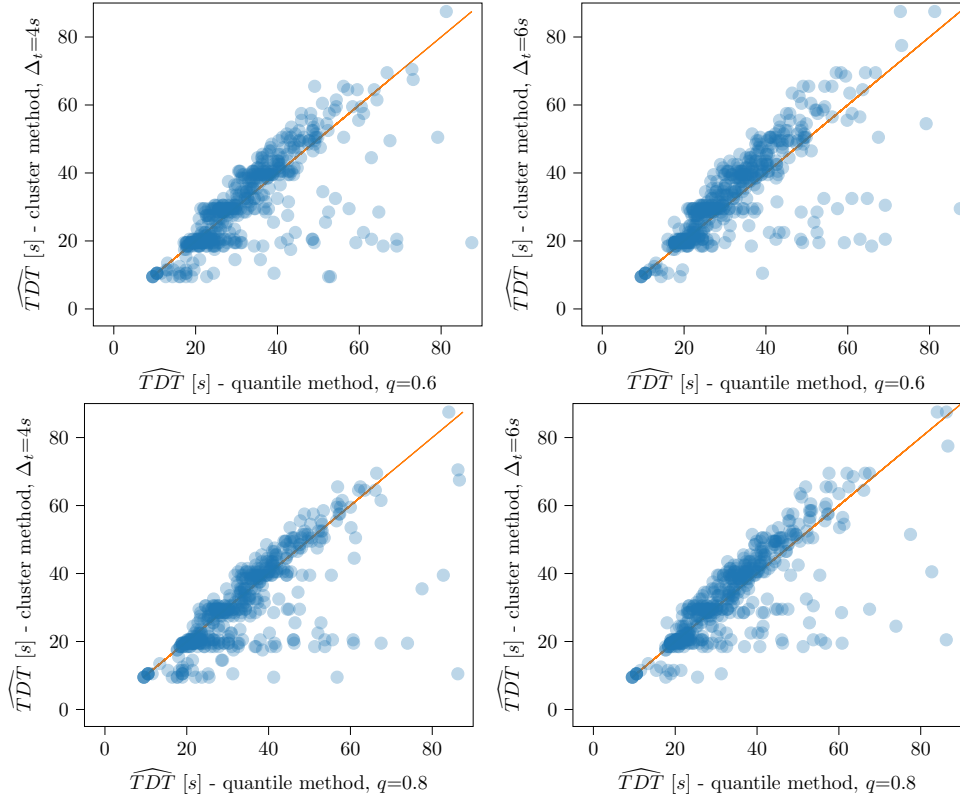
**Table 4** Subsets required to use the different methods.

584 the methods two by two where both are defined. Table 4 reminds the sub-sets on which  
585 each method is defined and highlights the subset used in the following subsections.

## 586 6.1 Quantile method vs Cluster method

587 First, we compare the quantile method and the cluster method. Figure 11 shows the  
588 comparison between these two methods with the sets of parameters ( $\Delta_t=4s$ ,  $\Delta_t=6s$ )  
589 and ( $q=0.6$ ,  $q=0.8$ ). We see a good correspondence between the two methods most of  
590 the time. The cases where the difference is important are cases of slow flows. This is  
591 the result of how passenger movements are taken into account in each method. The  
592 cluster method ignores the isolated passenger movements while the quantile method  
593 accounts for isolated movements. In practice, this means that if the flow is slow enough:  
594 the cluster method may ignore most of the passengers while the quantile method may  
595 add a particularly long time to the duration of the quantile. This results in a sizeable  
596 difference between the two estimations.

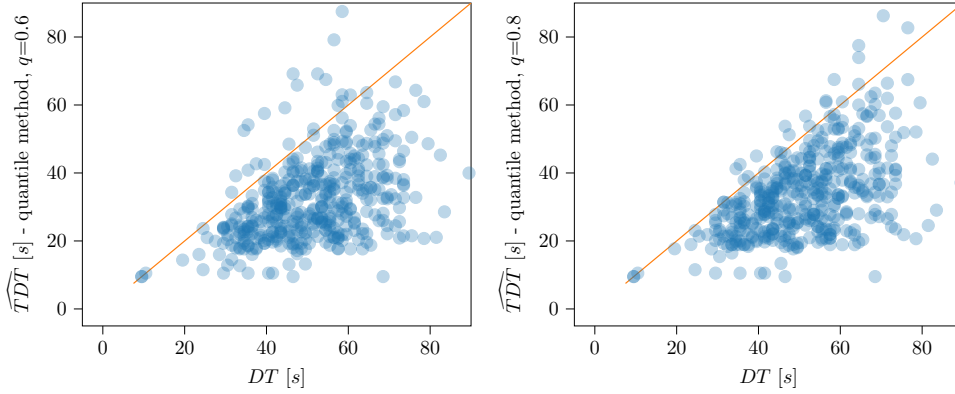
597 In the case of  $\Delta_t=6s$ , the cluster method gives values slightly greater than the  
598 quantile method.  $\Delta_t=6s$  often leads to considering all passengers in the main flow  
599 while the end of the alighting and boarding process is squeezed by the quantile method,  
600 this would mean  $\Delta_t=6s$  generally overestimate the tight dwell time.



**Figure 11** Comparison of estimated tight dwell times ( $\widehat{TDT}$ ) between the quantile method with  $q=0.6$  (top) and  $q=0.8$  (bottom) and cluster method with  $\Delta_t=4$  (left) and  $\Delta_t=6$  (right). Both methods give most of the time similar  $\widehat{TDT}$ . The points where the difference is significant may correspond to too early separation from the first cluster in the cluster method resulting in an under-estimation with this method.

601 The uncertainty on the estimated tight dwell time can be obtained from the uncer-  
602 tainty in data (both in time and passengers' movements). For passengers' movement,  
603 we assess the resulting uncertainty by assuming a uniform distribution of the error  
604 along the alighting and boarding process (without additional time). For the clus-  
605 ter method, simulations show 5.6s of root mean squared error on average for  $\Delta_t=4s$   
606 and 7.3s for  $\Delta_t=6s$ . For the quantile method, on average, there is no error due to  
607 misestimation of the number of alighting and boarding (as assumed uniform with  
608 time). However, one has to add uncertainty in counting event time that is at worst  
609 overestimated by 4.5s.

610 Without excluding the cluster method from practical estimations, the quantile  
611 method is more resilient in terms of data quality. For the sake of conciseness, we will  
612 compare the existing methods only with the quantile method in the following sections.



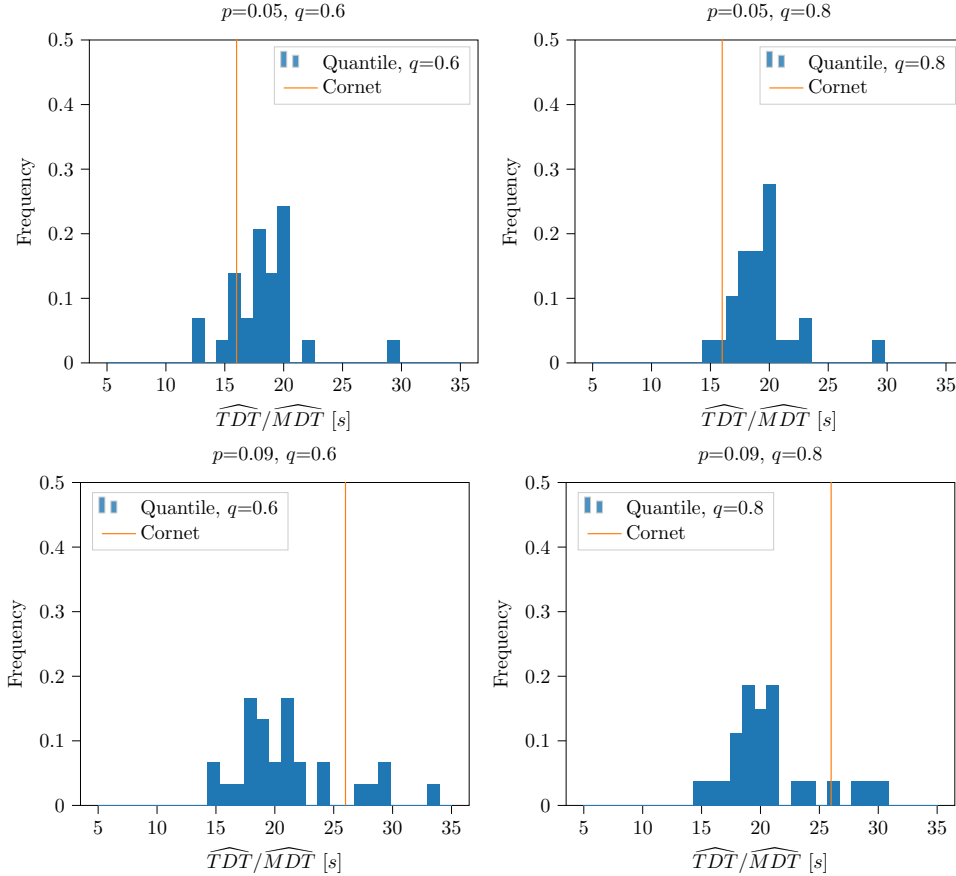
**Figure 12** Comparison of the actual observed dwell time ( $DT$ ) for late trains and the estimated tight dwell time ( $\widehat{TDT}$ ) with the quantile method for  $q=0.6$  (left) and  $q=0.8$  (right). Except for rare exceptions,  $DT$  is almost always larger than  $\widehat{TDT}$  for late trains.

## 6.2 Estimated tight dwell times with quantile method vs observed dwell times for late trains

In this subsection,  $\widehat{TDT}$  estimated using the quantile method (presented in Section 5.3.2) is compared to observed dwell times ( $DT$ ) for late trains (presented in Section 5.1). This comparison, shown in Figure 12 is performed using one station (*Versailles-Chantiers*) and one direction (toward Paris), and only trains for which the observed arrival time is greater than the scheduled departure time are selected. Observed dwell times for late trains are generally greater than  $\widehat{TDT}$  no matter their value. This suggests that the assumption that train drivers leave the station as soon as the alighting and boarding process is completed does not hold. Nevertheless, for a few points the computed  $\widehat{TDT}$  with the quantile method is greater than the actual  $DT$ . This is an artifact of the quantile method: if the chosen quantile  $q$  includes a high proportion of slower passengers, then  $\widehat{TDT}$  will be estimated as greater than the actual dwell time. Notice that for  $q=0.8$ , such points are rarer, which means that the extrapolation better fits the remaining isolated passengers in that case.

## 6.3 Estimated tight dwell times with quantile method vs Cornet's minimum dwell time

In this subsection, we compare  $\widehat{TDT}$  estimated with the quantile method with  $\widehat{MDT}$  estimated with Cornet's method. We only consider the stops where the observed departure time of the train is strictly greater than the scheduled departure time, which makes Cornet's method applicable. Figure 13 shows the histograms of  $\widehat{TDT}$  values for two values of  $p$  and two values of  $q$  ( $q=0.6$ ,  $q=0.8$ ) and the associated values of  $\widehat{MDT}$ . For those examples, the minimal values of  $\widehat{TDT}$  obtained with the quantile method are less than  $\widehat{MDT}$  obtained with Cornet's method. We can interpret this result as suggesting the existence of residual margins in Cornet's estimation. The spread

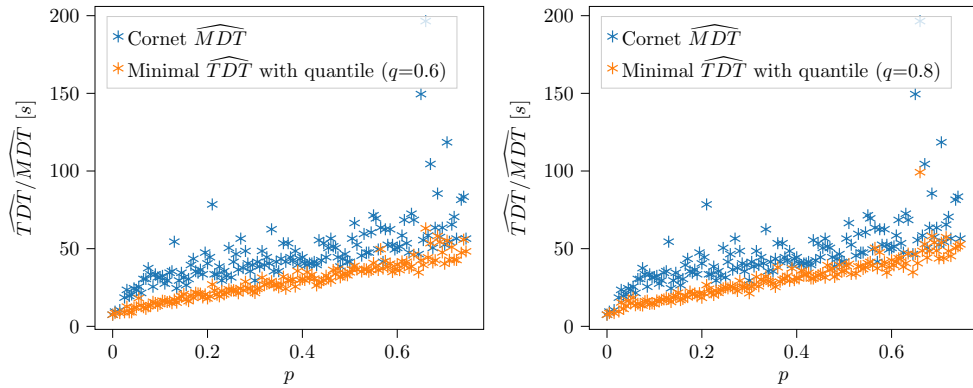


**Figure 13** Histogram of the tight dwell times ( $\widehat{TDT}$ ) estimated with the quantile method and the minimum dwell time ( $\widehat{MDT}$ ) estimated with Cornet's method for  $p=0.05$  (top) and  $p=0.09$  (bottom), for  $q=0.6$  (left) and  $q=0.8$  (right). Cornet's  $\widehat{MDT}$  is greater than the minimum of the quantile method's  $\widehat{TDT}$  which suggests potential residual margins in Cornet's estimation. For fixed  $p$ , the spread of  $\widehat{TDT}$  is large which means a more precise estimation with the quantile method.

638 of  $\widehat{TDT}$  values is non-negligible for one value of  $p$  while  $\widehat{MDT}$  is unique which means  
 639 that the quantile method can take into account the variability of the flow.

640 Figure 14 shows the  $\widehat{MDT}$  (estimated with Cornet's method) and minimal  $\widehat{TDT}$   
 641 (obtained with the quantile method) versus the reduced flow  $p$ . We cannot estimate  
 642  $\widehat{MDT}$  for high  $p$ , then the plotted values are for  $p \leq 0.7$ . The values obtained with the  
 643 quantile method are always equal or less than Cornet's  $\widehat{MDT}$ . As already mentioned,  
 644 nothing can guarantee Cornet's  $\widehat{MDT}$  to be free of margin while the tight dwell time  
 645 obtained with the quantile method is defined as free of margin. Moreover, a quantile  
 646 of the flow is not the total flow: for instance, if people in the quantile are much faster  
 647 than the others, this will lead to an underestimation of  $\widehat{TDT}$  (and vice versa). Hence,





**Figure 14** Comparison of Cornet minimum dwell times ( $\widehat{MDT}$ ) and minimal tight dwell time ( $\widehat{TDT}$ ) obtained with the quantile method versus the reduced flow  $p$  for  $q=0.6$  (left) and  $q=0.8$  (right). In both cases, Cornet's method gives greater  $\widehat{MDT}$  than the quantile's minimal  $\widehat{TDT}$  for all values of  $p$  which highlights residual margins in Cornet's method.

648 the minimal value of  $\widehat{TDT}$  for a given  $p$  can be obtained for a particularly fast quantile  
 649 compared to the rest of the flow which could concur in explaining the difference.

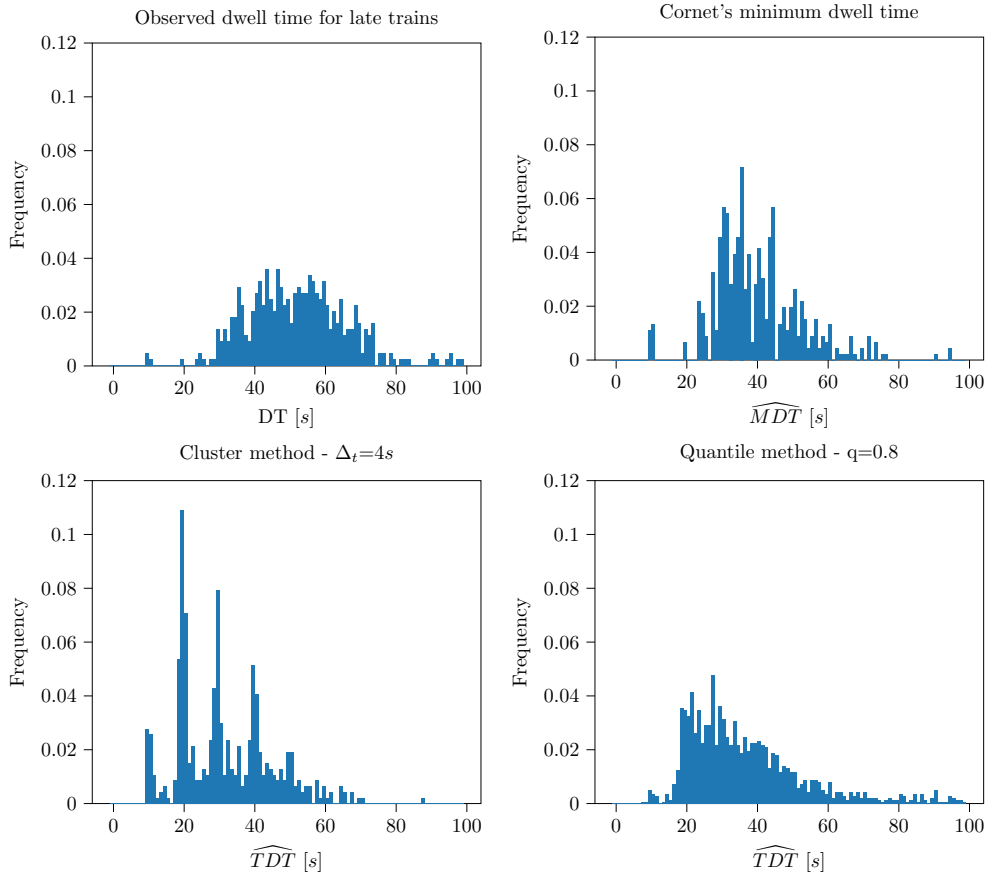
#### 650 6.4 Summary of the comparisons

651 In this subsection, we summarize the previous results. We review the distribution of  
 652 the value obtained with each method, fixing one appropriate parameter for the cluster  
 653 and the quantile method is required to do so. We chose  $\Delta_t=4s$  for the cluster method  
 654 as done by Wiggendaad (2001) (for their comparable manual method) while  $q=0.8$   
 655 is selected for the quantile method. Once again any  $\Delta_t \geq 4s$  or any  $0.5|q|1$  should  
 656 give similar results. To make everything comparable, we consider only late trains  
 657 (those having their observed arrival time greater than their scheduled departure time).  
 658 After that, we assess the methods based on additional qualitative indicators. These  
 659 indicators evaluate how relevant each method is for practical use.

660 First, Figure 15 shows the statistical distribution of the values obtained with each  
 661 method. We see that all the methods give values generally smaller than the actual  
 662 observed dwell time  $DT$ . We can also observe that Cornet's  $\widehat{MDT}$  is similar to the  
 663 values of both the cluster and the quantile method but with a less important range  
 664 between small and high values. The quantile method gives a  $\widehat{TDT}$  that is slightly  
 665 greater than the cluster method. It is worth pointing out the peaks around  $\widehat{TDT}=20s$ ,  
 666  $\widehat{TDT}=30s$  and  $\widehat{TDT}=40s$  with the cluster method. These peaks are directly correlated  
 667 with the peak at 10s in the distribution of intervals between two counting events  
 668 (Figure 5 in Section 4.1).

669 We are also interested in the practical relevance of the methods described here,  
 670 in addition to comparing the outcome of the different methods. Table 5 qualitatively  
 671 sums up the pros and cons of each method over the following criterion:

- 672 • The robustness relative to data lack of quality,
- 673 • The resolution of the data used (see Section 3.1),



**Figure 15** Distribution of the results of the four presented methods for the subset of trains arriving later than their theoretical departure time (467 observations).  $\widehat{TDT}$  and  $\widehat{MDT}$  are generally smaller than  $DT$ . The cluster and quantile method give  $\widehat{TDT}$  slightly smaller than Cornet's  $\widehat{MDT}$  but with a higher dispersion.

- 674 • The range of usability if the data is accessible,
- 675 • The amount of necessary data to perform one estimation.

676 The quantile method has both a high robustness and resolution compared to the  
 677 other methods described in this paper; it has a large range of usability when the data  
 678 is accessible. The amount and quality of the necessary data is the only drawback of  
 679 the quantile method compared to the other methods described here. Nevertheless, the  
 680 data required is provided continuously by the APC system which makes it easy to  
 681 access once the system is available. Furthermore, as the APC technology improves and  
 682 becomes more widespread, access to data will become less of an issue in the future.  
 683 Based on this we can conclude that the quantile method is the most appropriate way  
 684 to estimate a tight dwell time from a practical standpoint.

Criterion	Observed dwell time for late trains	Cornet's minimum dwell time	Cluster method	Quantile method
Robustness	++	+	-	++
Resolution	--	-	++	++
Range of usability	-	--	+	+
Amount of necessary data	+	-	-	-

**Table 5** Summary of qualitative comparisons of the four presented methods. The quantile method is the most appropriate way to estimate a tight dwell time for practical implications.

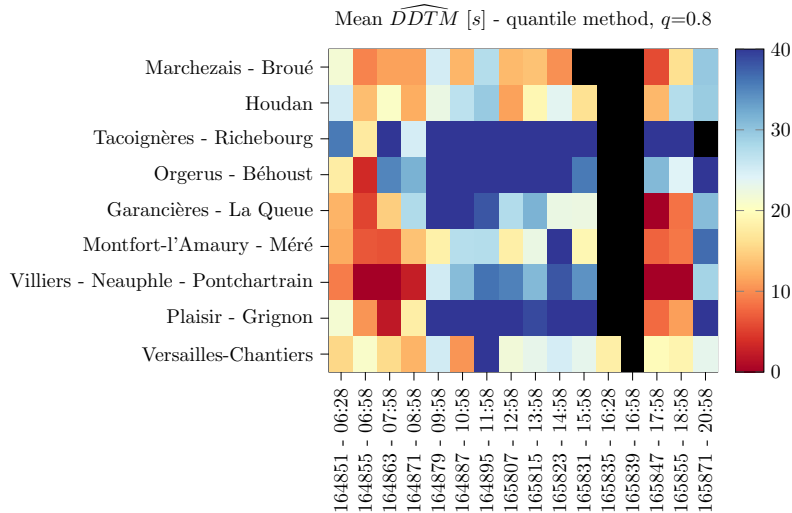
## 685 7 Use case on the line N

686 As mentioned earlier, the main implication of the tight dwell time is to estimate a  
687 posteriori the dwell time margins (subtracting the tight dwell time from the observed  
688 dwell time). In this section, we apply the quantile method to estimate the dwell time  
689 margins for all stations of the branch Paris-Dreux on *Line N* (presented in Section  
690 4.1) on weekdays. The terminus (*Paris-Montparnasse* and *Dreux*) were removed since  
691 there is no point in estimating dwell time margins at a terminus. We selected the  
692 couples  $(k,s)$  where we had at least ten dates of observations ( $d$ ) so we could average  
693 over a decent sample. This resulted in the absence of certain train numbers ( $k$ ) in our  
694 analysis due to gaps in the data (see Section 4.1). Figure 16 shows the average values  
695 of  $\widehat{DTM}$  for all the stations and all the services of this branch toward Dreux computed  
696 with the quantile method ( $q=0.8$ ) between September 2022 and March 2023. The first  
697 observation that could be made is that substantial dwell time margins are observed  
698 in many cases (around 20s on average). The main stations of the line (*Versailles-*  
699 *Chantiers* and *Plaisir - Grignon*) handle a larger volume of passengers but also have  
700 a large margin on average. This is likely due to longer scheduled dwell times at these  
701 stations. Some services also had lower  $\widehat{DTM}$  on average (particularly visible for the  
702 two main stations). These services corresponded to the evening peak hours. *Line N*  
703 being a commuter train, the direction toward Dreux was the busiest during the evening  
704 peak hours. The additional scheduled dwell time at *Versailles-Chantiers* and *Plaisir -*  
705 *Grignon* is consumed during the evening peak hours. One can notice that the station  
706 *Houdan* had quite low  $\widehat{DTM}$  on average, whereas it was not a connection station nor  
707 did it gather some specific point of interest. So far, we have no explanation for these  
708 values.

709 Figure 17 shows the same illustration for the opposite direction (toward Paris).  
710 Similar observations can be made except for *Montfort-l'Amaury - Méré* where sub-  
711 stantial dwell time margins can be observed. This effect is likely due to the convergence  
712 after *Villiers - Neauphle - Pontchartrain*. Additional time is scheduled to make sure  
713 the train path is respected before approaching the convergence.

714 We also observe higher dwell time margins in the evening peak hours toward Dreux  
715 than in the morning peak hours toward Paris. That could be explained by a more  
716 homogeneous distribution of the passengers' movements in the evening than in the  
717 morning.

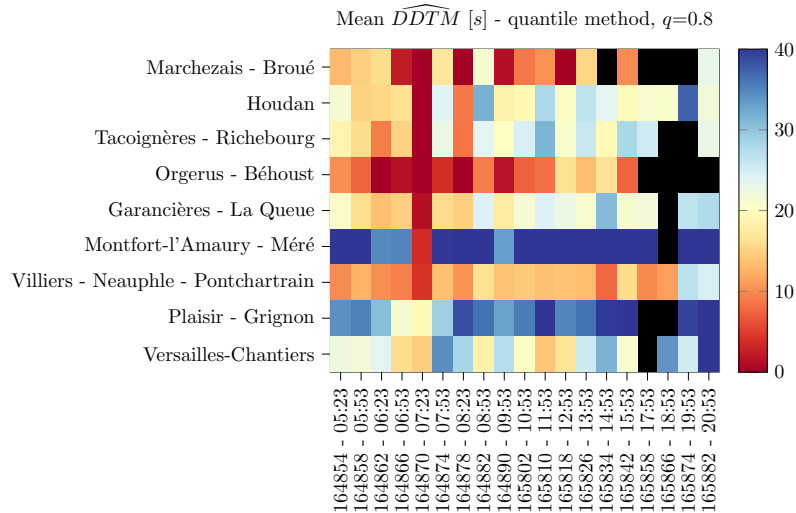
718 This assumption is partly validated by Figure 18 which shows the mean door dwell  
719 time margin  $\widehat{DDTM}$  estimated for the Paris-Dreux services at *Versailles-Chantiers*.



**Figure 16** Average estimated dwell time margins ( $\widehat{DTM}$ ) for all services on line N, branch Paris-Dreux toward Dreux computed with the quantile method,  $q=0.8$ . Black corresponds to a lack of data. The lowest  $\widehat{DTM}$  are obtained for the connection stations in the evening peak hours which is coherent with the commuting nature of line N.

720 We find that the front doors are often the busiest toward Paris with some doors with  
721 a particularly short  $\widehat{DDTM}$  (door 2 for service 164866 or door 6 for service 164870).  
722 Toward Dreux, the mean  $\widehat{DDTM}$  values are more homogeneous. For services 165855  
723 in particular, the minimal  $\widehat{DDTM}$  is not that low on average with very little difference  
724 from one door to another. Toward Paris, doors 1 to 8 are almost always busier than  
725 doors 9 to 16, this is probably because the terminus *Paris-Montparnasse* is a dead-end  
726 station where the station exit is at the front of the train. The same explanation may  
727 hold for trains toward Dreux where the pattern is the opposite where doors 9 to 16 are  
728 sometimes busier, probably because they are closer to the entrance of the platform.

729 Being busy does not necessarily mean many people are involved in the alighting  
730 and boarding process. A widespread idea is that alighting and boarding times at the  
731 door level are linear with passengers' movement. Figure 19 mitigates this general idea,  
732 showing the estimated  $\widehat{DABT}$  with the quantile method ( $q=0.8$ ) versus the number  
733 of passenger movements at *Versailles-Chantiers*, toward Paris, at door 8. Despite an  
734 increasing trend, we see a large spread of the  $\widehat{DABT}$  for a given number of passenger  
735 movements. This is due to the difference between alighting and boarding but also to the  
736 variability of the flows regarding passengers' behavior (e.g. more stressed passengers  
737 moving faster or boarding passengers in front of the doors). The attempted linear fit  
738 is of poor quality (having  $R^2=0.44$ ) and it does not cross the origin. This is probably  
739 due to a change of regime for more than ten passengers' movements:  $\widehat{DABT}$  being  
740 more spread and increasing less fast for more than ten passengers' movements. Such  
741 a change could be due to crowding effects that will be studied in future works.



**Figure 17** Average estimated dwell time margins ( $\widehat{DTM}$ ) for all services on line N, branch Paris-Dreux toward Paris computed with the quantile method,  $q=0.8$ . Black corresponds to a lack of data. The lowest  $\widehat{DTM}$  are obtained for the connection stations in the morning peak hours which is coherent with the commuting nature of line N.

## 742 8 Conclusions and perspectives

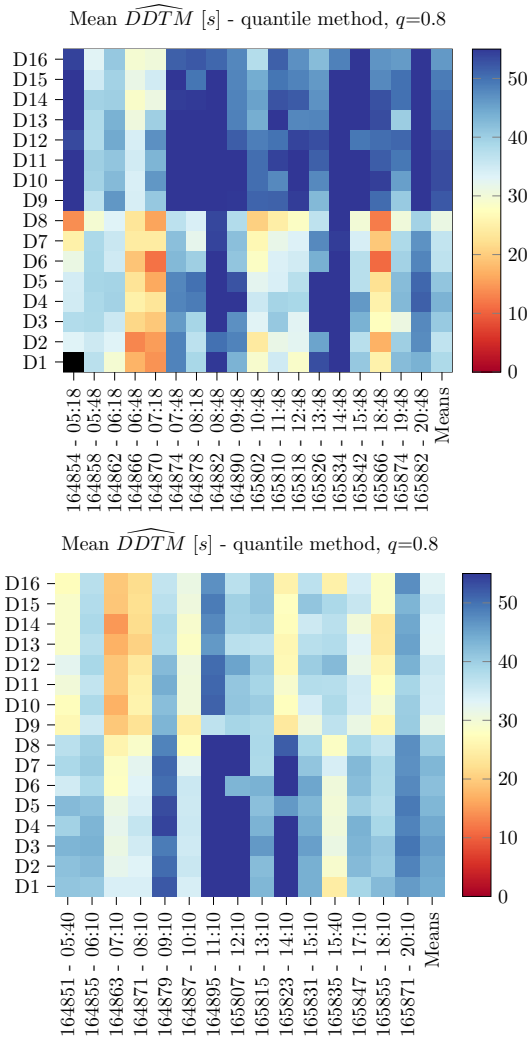
### 743 8.1 Main research findings

744 New passenger flow data with a higher resolution being accessible, we developed  
 745 two new methods enabling us to compute the tight dwell time (dwell time without  
 746 margins).

747 Contrary to existing methods to estimate tight dwell times that are only applicable  
 748 when using data from stops of late trains, the cluster method and the quantile method  
 749 presented in this paper make it possible to estimate the tight dwell times for any  
 750 stop and any train. In addition to this, the methods allow us us us us to estimate  
 751 the alighting and boarding time for any door. Compared to the quantile method, the  
 752 existing heuristics show weaknesses. Taking the observed dwell times of late trains is  
 753 an overestimation of the tight dwell time while Cornet's minimum dwell time does  
 754 not take into account enough variability. For the same number of passengers alighting  
 755 and boarding, the necessary time to complete the alighting and boarding process may  
 756 vary and change the tight dwell time. Our proposed methods capture this variability.

### 757 8.2 Implications for practice

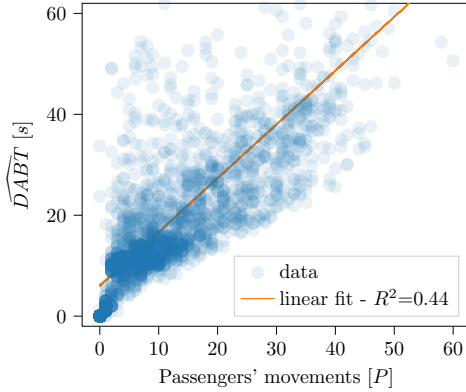
758 The extension of the estimation of tight dwell times to any stop enables the compu-  
 759 tation of dwell time margins for past circulations. Knowing those values will help to  
 760 make informed decisions to optimize the timetable for the future. Having an estimation  
 761 for the alighting and boarding time at all the doors of the train gives an illustration of



**Figure 18** Average estimated door dwell time margins ( $\widehat{DDTM}$ ) for all doors and all services on line N, branch Paris-Dreux at *Versailles-Chantiers* toward Paris (top), toward Dreux (bottom) computed with the quantile method,  $q=0.8$ . Black corresponds to a lack of data. Toward Paris, the busiest doors are the ones in the front of the train while it is the opposite toward Dreux; in coherence with Paris being a dead-end station.

762 the distribution of the passengers' movements in the platform train interface. For ex-  
 763 ample, some doors may have low margins while others may have higher ones. Such an  
 764 illustration would help to evaluate policies to make the alighting and boarding times  
 765 more uniform among the doors.

766 For lines or networks that would not be provided high-resolution passenger  
 767 counting data, the above-presented results show quite close values between Cornet's



**Figure 19** Door alighting and boarding times ( $\widehat{DABT}$ ) at door 8 for all services on *Line N* at *Versailles-Chantiers* toward Paris versus the number of passenger movements computed with the quantile method,  $q=0.8$ . Despite an increasing trend,  $\widehat{DABT}$  is not well modeled by a linear fit of passengers' movements.

768 minimum dwell times and tight dwell times. Only stop-scale counting data (more eas-  
 769 ily available thanks to smartcard data or video counting) may be enough to build  
 770 decent timetables.

### 771 8.3 Limitations of the current study

772 In this study, the assessment of the new methods we proposed was done by comparison  
 773 to existing methods. To ensure a fair comparison, we made the estimations in similar  
 774 subsets of the data (i.e. for late trains or non-early trains, see Table 4). This means  
 775 we have no direct assessment for non-late trains and the extrapolation of the results  
 776 for late trains should be done cautiously. Furthermore, the comparison we made was  
 777 relative to the existing methods. We highlighted that Cornet's minimum dwell time  
 778 keeps little margin but it is possible that our methods may also have underestimated  
 779 the tight dwell time (depending on their parameter  $\Delta_t$  or  $q$ ). It is, however, non-  
 780 trivial to measure precisely to what extent. The main issue there was that we did not  
 781 compare the same object. It is possible that in a comparable situation,,,,, we could have  
 782 observed people alighting and boarding faster which would have given a smaller  $TDT$ .  
 783 The methods we proposed compute the  $TDT$  as  $TDT = ABT + TT$  ( $TT$  being the  
 784 technical time). This suggests that everything that is not included in the alighting and  
 785 boarding time is margin and should be dropped to compute  $TDT$ . This also implies  
 786 that the time for a particularly slow passenger to alight/board the train should be  
 787 considered as some margin (slow regarding the parameter  $\Delta_t$  or  $q$ ) and we know that  
 788 it is not always the case. Such an issue could be discussed in more detail using video  
 789 of the dwelling process but we only have access to passenger counting data.

### 790 8.4 Future research directions

791 As already mentioned, the methods we developed enabled us to analyze the tight dwell  
 792 times for past circulations. Future research should focus on predicting  $TDT$  for future

793 circulations. Knowing obtained  $\widehat{TDT}$ , some dwell time margin will have to be smartly  
794 added to build a robust timetable for an existing line. Some more advanced work will  
795 be able to make possible predictions of  $TDT$  for new missions, new stops, and new  
796 lines.

797 The study presented here is limited to discussing two novel methods to estimate  
798 tight dwell times using high resolution passenger count data. Future studies could  
799 make use of other observations methods, such as the use of video footage, to determine  
800 the accuracy of the methods proposed here. This fell outside of the scope of the study  
801 presented here.

802 Finally, the study proposed here only discussed one implication of high-resolution  
803 passenger counting data: estimation of tight dwell times and dwell time margins.  
804 Future work will explore other implications for better understanding the alighting and  
805 boarding process.

## 806 Compliance with Ethical Standards

807 **Conflict of Interest.** We declare to have no financial and non-financial conflict of  
808 interests or any interests that might be perceived to influence neither the results nor  
809 the discussion reported in this paper.

810 **Funding.** This work was supported by the Association Nationale de la Recherche et  
811 de la Technologie (ANRT) (Grant numbers 2022/1049). The data used was provided  
812 by SNCF Voyageurs. The authors declare that no other funds, grants, or other support  
813 were received during the preparation of this manuscript.

## 814 References

- 815 Andersson, E., Peterson, A., & Krasemann Törnquist, J. (2011). Robustness in swedish  
816 railway traffic timetables. In *4th international seminar on railway operations modelling  
817 and analysis*. Rome, Italy.
- 818 Andersson, E., Peterson, A., & Törnquist Krasemann, J. (2013). Quantifying rail-  
819 way timetable robustness in critical points. *Journal of Rail Transport Planning &  
820 Management*, 3(3), 95-110. Retrieved from [https://www.sciencedirect.com/science/  
821 article/pii/S2210970613000516](https://www.sciencedirect.com/science/article/pii/S2210970613000516) (Robust Rescheduling and Capacity Use) doi:  
822 doi:10.1016/j.jrtpm.2013.12.002
- 823 Armstrong, J., & Preston, J. (2011). Alternative railway futures: growth  
824 and/or specialisation? *Journal of Transport Geography*, 19(6), 1570-1579. Re-  
825 trieved from <https://www.sciencedirect.com/science/article/pii/S0966692311001281> doi:  
826 doi:10.1016/j.jtrangeo.2011.03.012
- 827 Brethomé, L. (2018). *Modélisation et optimisation d'un plan de transport ferroviaire en zone  
828 dense du point de vue des voyageurs* (Doctoral thesis, École Centrale de Lille). Retrieved  
829 from <https://www.theses.fr/2018ECLI0014>
- 830 Buchmueller, S., Weidmann, U., & Nash, A. (2008). Development of a dwell time calculation  
831 model for timetable planning. *WIT Transactions on the Built Environment*, 103, 525-534.  
832 doi: doi:10.2495/CR080511
- 833 Christoforou, Z., Chandakas, E., & Kaparias, I. (2020). Investigating the impact of dwell  
834 time on the reliability of urban light rail operations. *Urban Rail Transit*, 6, 116-131. doi:  
835 doi:10.1007/s40864-020-00128-1



- 836 Cornet, S., Buisson, C., Ramond, F., Bouvarel, P., & Rodriguez, J. (2019). Methods for  
837 quantitative assessment of passenger flow influence on train dwell time in dense traf-  
838 fic areas. *Transportation Research Part C: Emerging Technologies*, 106, 345-359. doi:  
839 doi:10.1016/j.trc.2019.05.008
- 840 Coulaud, R., Keribin, C., & Stoltz, G. (2023). Modeling dwell time in a data-rich railway  
841 environment: with operations and passenger flows data. *Transportation Research Part C:  
842 Emerging Technologies*, 146, 103980. doi: doi:10.1016/j.trc.2022.103980
- 843 Daamen, W., Goverde, R. M., & Hansen, I. A. (2009). Non-discriminatory automatic  
844 registration of knock-on train delays. *Networks and Spatial Economics*, 9, 47-61.
- 845 De Nailly, P. (2023). *Analyse et prédiction des flux piétons dans un pôle de transport*  
846 *multimodal à partir de données multi-sources* (Doctoral thesis, Université Gustave Eiffel).  
847 Retrieved from <https://theses.hal.science/tel-04090167>
- 848 Dewilde, T., Sels, P., Cattrysse, D., & Vansteenwegen, P. (2011). Defining robustness of  
849 a railway timetable. In *4th international seminar on railway operations modelling and*  
850 *analysis*. Rome, Italy.
- 851 DRIEA, O., Ile-de-France Mobilité. (2022). *Enquête globale transport (egt) - 2020*.
- 852 Givoni, M., Brand, C., & Watkiss, P. (2009). Are railways climate friendly? *Built Environ-*  
853 *ment*, 35(1), 70-86. Retrieved from [https://www.ingentaconnect.com/content/alex/benv/  
854 2009/00000035/00000001/art00006](https://www.ingentaconnect.com/content/alex/benv/2009/00000035/00000001/art00006) doi: doi:10.2148/benv.35.1.70
- 855 Goverde, R. (2005). *Punctuality of railway operations and timetable stability analysis*  
856 (Doctoral thesis, TU Delft). Retrieved from [https://www.researchgate.net/profile/  
857 Rob-Goverde/publication/27346099\\_Punctuality\\_of\\_railway\\_operations\\_and\\_timetable  
858 \\_stability\\_analysis/links/00463528a2e4892917000000/Punctuality-of-railway-operations  
859 -and-timetable-stability-analysis.pdf](https://www.researchgate.net/profile/Rob-Goverde/publication/27346099_Punctuality_of_railway_operations_and_timetable_stability_analysis/links/00463528a2e4892917000000/Punctuality-of-railway-operations-and-timetable-stability-analysis.pdf)
- 860 Hansen, I. (2010). Railway network timetabling and dynamic traffic management. *Interna-*  
861 *tional Journal of Civil Engineering*, 8. Retrieved from [http://ijce.iust.ac.ir/article-1-422  
862 -en.html](http://ijce.iust.ac.ir/article-1-422-en.html)
- 863 Harris, N. (2005). Train boarding and alighting rates at high passenger loads. *Journal of*  
864 *Advanced Transportation*, 40(3), 249-263. doi: doi:10.1002/atr.5670400302
- 865 Harris, N. (2015). A european comparison of station stop delays. In *International congress*  
866 *on* (p. 1-6).
- 867 IenW. (2022). *Programma hoogfrequent spoorvervoer - achtste oortgangsrapportage*. Re-  
868 trieved from [https://www.rijksoverheid.nl/documenten/rapporten/2022/10/11/bijlage-1  
869 -phs-vgr-7-2022-1](https://www.rijksoverheid.nl/documenten/rapporten/2022/10/11/bijlage-1-phs-vgr-7-2022-1)
- 870 Kecman, P., & Goverde, R. M. (2013). Process mining of train describer event data and  
871 automatic conflict identification. *Computers in Railways XIII: Computer System Design  
872 and Operation in the Railway and Other Transit Systems*, 127, 227.
- 873 Kecman, P., & Goverde, R. M. (2015). Predictive modelling of running and dwell times in  
874 railway traffic. *Public Transport*, 7, 295-319. doi: doi:10.1007/s12469-015-0106-7
- 875 Kuipers, R. (2024). Understanding dwell times using automatic passenger count data: A  
876 quantile regression approach. *Journal of Rail Transport Planning & Management*, 29,  
877 100431. doi: doi:10.1016/j.jrtpm.2024.100431
- 878 Kuipers, R., & Palmqvist, C.-W. (2022). Passenger volumes and dwell times for commuter  
879 trains: A case study using automatic passenger count data in stockholm. *Applied Sciences*,  
880 12. doi: doi:10.3390/app12125983
- 881 Li, D., Daamen, W., & Goverde, R. M. (2016). Estimation of train dwell time at short stops  
882 based on track occupation event data: A study at a dutch railway station. *Journal of*  
883 *Advanced Transportation*, 50(5), 877-896.
- 884 Luan, X., & Corman, F. (2022). Passenger-oriented traffic control for rail networks:  
885 An optimization model considering crowding effects on passenger choices and train

- 886 operations. *Transportation Research Part B: Methodological*, 158, 239-272. Re-  
887 trieved from <https://www.sciencedirect.com/science/article/pii/S0191261522000273> doi:  
888 doi:10.1016/j.trb.2022.02.008
- 889 Matthew Volovski, C. C., Evangelia S. Ieronymaki, & O'Loughlin, J. P. (2021). Subway sta-  
890 tion dwell time prediction and user-induced delay. *Transportmetrica A: Transport Science*,  
891 17(4), 521–539. doi: doi:10.1080/23249935.2020.1798555
- 892 McNally, M. G. (2007). The four step model. In *Handbook of transport modeling* (chap. 3).  
893 Medeossi, G., & Nash, A. (2020). Reducing delays on high-density railway lines:  
894 London–shenfield case study. *Transportation Research Record*, 2674(7), 193–205. doi:  
895 doi:10.1177/0361198120921159
- 896 Mouwen, A., & Rietveld, P. (2013). Does competitive tendering improve customer satisfaction  
897 with public transport? a case study for the netherlands. *Transportation Research Part A:  
898 Policy and Practice*, 51, 29-45. doi: doi:10.1016/j.tra.2013.03.002
- 899 Palmqvist, C.-W. (2019). *Delays and timetabling for passenger trains* (Doctoral thesis, Lund  
900 University Faculty of Engineering, Technology and Society, Transport and Roads). Re-  
901 trieved from [http://portal.research.lu.se/ws/files/70626078/Carl\\_William\\_Palmqvist\\_web](http://portal.research.lu.se/ws/files/70626078/Carl_William_Palmqvist_web.pdf)  
902 [.pdf](http://portal.research.lu.se/ws/files/70626078/Carl_William_Palmqvist_web.pdf)
- 903 Pedersen, T., Nygreen, T., & Lindfeldt, A. (2018). Analysis of temporal factors influencing  
904 minimum dwell time distributions. *WIT Transactions on the Built Environment*, 181,  
905 447–58. doi: doi:10.2495/CR180401
- 906 Puong, A. (2000). Dwell time model and analysis for the mbta red line. *Massachusetts  
907 Institute of Technology Research Memo*, 02139–4307.
- 908 Seriani, S., & Fujiyama, T. (2019). Modelling the distribution of passengers waiting to  
909 board the train at metro stations. *Journal of Rail Transport Planning & Management*,  
910 11, 100141. doi: doi:10.1016/j.jrtpm.2019.100141
- 911 van Loon, R., Rietveld, P., & Brons, M. (2011). Travel-time reliability impacts on  
912 railway passenger demand: a revealed preference analysis. *Journal of Transport Geogra-  
913 phy*, 19(4), 917-925. Retrieved from [https://www.sciencedirect.com/science/article/pii/](https://www.sciencedirect.com/science/article/pii/S0966692310001912)  
914 [S0966692310001912](https://www.sciencedirect.com/science/article/pii/S0966692310001912) doi: doi:10.1016/j.jtrangeo.2010.11.009
- 915 Vromans, M. (2005). *Reliability of railway systems* (Doctoral thesis, Erasmus Universiteit  
916 Rotterdam). Retrieved from <https://repub.eur.nl/pub/6773/>
- 917 Weston, J. (1989). Train service model-technical guide. *London underground operational  
918 research note*, 89, 18.
- 919 Wiggeraad, I. P. B. (2001). *Alighting and boarding times of passengers at dutch railway  
920 stations* (Doctoral thesis). TU Delft.
- 921 Yamamura, A., Koresawa, M., Adachi, S., & Tomii, N. (2012). Identification of causes of  
922 delays in urban railways. *WIT Transactions on The Built Environment*, 127, 403–414.  
923 doi: doi:10.2495/CR120341
- 924 Yang, J., Shiwakoti, N., & Tay, R. (2019). Train dwell time models – development in the past  
925 forty years. In *41st australasian transport research forum (atrf)*. Canberra, Australia.
- 926 Yap, M., & Cats, O. (2021). Predicting disruptions and their passenger delay impacts for  
927 public transport stops. *Transportation*, 48, 1703–1731. doi: doi:10.1007/s11116-020-10109-  
928 9

929 **9 Appendix**

930 **9.1 Demonstration of the comparability of Cornet's  $\widehat{MDT}$  and**  
 931 **estimated  $\widehat{TDT}$**

932 In this section, we demonstrate that if  $\widehat{ABT}$  and  $\widehat{DTM}$  are independent and

933  $\lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{DTM}_{k,s_0,d} \right) = 0$ , then:

934 
$$\lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{TDT}_{k,s_0,d} \right) = \lim_{|\{k,d\}| \rightarrow \infty} \widehat{MDT}(p_0, s_0).$$

935 First, we trivially have (without any hypothesis):

936 
$$\min_{k,d \text{ s.t. } p=p_0} \widehat{TDT}_{k,s_0,d} \leq \widehat{MDT}(p_0, s_0),$$

937 which holds if one takes the limit.

938 Second, for any  $\epsilon > 0$ , there exists  $n$  a positive integer, such that if  $|\{k,d\}| \geq n$ ,  
 939 there exists  $(k_0, d_0)$  such that:

940 
$$\begin{cases} \widehat{ABT}_{k_0,s_0,d_0} \leq \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{ABT}_{k,s_0,d} \right) + \frac{\epsilon}{2} \\ \widehat{DTM}_{k_0,s_0,d_0} \leq \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{DTM}_{k,s_0,d} \right) + \frac{\epsilon}{2}, \end{cases}$$

941 as the estimation of  $\widehat{ABT}_{k,s,d}$  and  $\widehat{DTM}_{k,s,d}$  is supposed independent. Assuming that:

942 
$$\lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{DTM}_{k,s_0,d} \right) = 0.$$

943 This implies that:

944 
$$\widehat{ABT}_{k_0,s_0,d_0} + \widehat{DTM}_{k_0,s_0,d_0} \leq \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{ABT}_{k,s_0,d} \right) + \epsilon.$$

945 In particular,

946 
$$\left( \min_{k,d \text{ s.t. } p=p_0} \widehat{ABT}_{k,s_0,d} + \widehat{DTM}_{k,s_0,d} \right) \leq \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{ABT}_{k,s_0,d} \right) + \epsilon.$$

947 This being true for any  $\epsilon$ , one obtains:

948 
$$\lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{ABT}_{k,s_0,d} + \widehat{DTM}_{k,s_0,d} \right) \leq \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{ABT}_{k,s_0,d} \right).$$

949 And, adding  $TT$ ,

$$950 \quad \lim_{|\{k,d\}| \rightarrow \infty} \widehat{MDT}(p_0, s_0) \leq \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{TDT}_{k,s_0,d} \right).$$

951 Therefore,

$$952 \quad \lim_{|\{k,d\}| \rightarrow \infty} \widehat{MDT}(p_0, s_0) = \lim_{|\{k,d\}| \rightarrow \infty} \left( \min_{k,d \text{ s.t. } p=p_0} \widehat{TDT}_{k,s_0,d} \right).$$

953 In practice, if we see that the minimum in a given window of  $p$  of  $\widehat{TDT}$  we measure  
954 with our methods is close enough to  $\widehat{MDT}$  estimated for the same value of  $p$ , this  
955 means that the methods are coherent. By contrast, if the minimum in a given window  
956 of  $p$  of  $\widehat{TDT}$  we measure with our methods is different from  $\widehat{MDT}$  for the same  $p$ ,  
957 then either the minimal dwell time margin is not zero, or the amount of data is not  
958 large enough, or one of the methods is not coherent.