¹ Estimation of train dwell times based on high-resolution passenger flow data.

| Mehdi Baali ¹ , Ruben Kuipers ² , Rémi Coulaud ³ , Christine Buisson ⁴ , and |
|--|
| $Carl-William Palmqvist^5$ |
| |

¹Ph.D. Student, SNCF-Transilien, LICIT-Éco7, Université Gustave Eiffel, France
 ²Ph.D. Student, Department of Technology and Society, Lund University, Sweden
 ³Head of Datalab' MT, SNCF-Transilien, France
 ⁴Senior researcher, LICIT-Éco7, Université Gustave Eiffel, France
 ⁵Assistant professor, Department of Technology and Society, Lund University, Sweden

Abstract

Dwell time is a crucial stake for railway operations corresponding to 20% of the 10 total travel time in a mass transit context. A relevant margin must be allocated to the 11 dwell time to guarantee the robustness of a timetable. Several studies have evaluated 12 the robustness of the running time. However, except for a few heuristics, little work has 13 tackled the issue of measuring the dwell time without margin named tight dwell time 14 in this work. Then, we present two methods to estimate the tight dwell time from high-15 resolution passenger flow data. Given the access to this data, both methods enable 16 estimating the tight dwell time for all the stops while existing heuristics are limited 17 to late trains and/or few passengers. Besides, our developed methods highlight the 18 propensity of existing heuristics to overestimate what they measure. The estimation 19 ex-post of the tight dwell time would help the design of future timetables. 20

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

23 1 INTRODUCTION

23

4

5

6

7

8

9

The last decade has seen a change in the way the transport sector is viewed, with the 24 sustainability of the system being called into question (Armstrong & Preston, 2011). In 25 line with this, efforts are being made to increase the modal share of more sustainable 26 modes of transport and induce a shift away from private motorized transport. Railways 27 are one of the modes that can play a vital role in this process and make transportation 28 systems more sustainable from an environmental point of view. However, despite railways 29 providing a green alternative to road transport the potential environmental benefits will 30 only come to fruition when ridership is sufficiently high (Givoni et al., 2009), meaning that 31 it is important to make railway an attractive mode of transport. The attractiveness of 32 railways is affected by the quality of service that is offered such as the service frequency 33 and on-time performance of railways, along with travel speed and vehicle tidiness (Mouwen 34 & Rietveld, 2013; van Loon et al., 2011). These quality aspects of railways present a trade-35 off, however, since an emphasis on the frequency of services often comes at the cost of 36 on-time performance (Mouwen & Rietveld, 2013). To ensure the quality of railways it is 37 thus important to balance both the frequency as well as the on-time performance. One 38 way in which railway operators can directly influence this balance is through the design 39 of a timetable. The timetable forms the backbone of railway systems by providing the 40 location and direction of trains at specified points in time (Goverde, 2005) and indicates 41 the service offered to passengers, showing both when and how often a train will run. 42

43 1.1 Timetabling

Timetable planning, or timetabling, is the task of scheduling when and where trains will run within a given network, matching the desired train routes to the available infrastructure

(Goverde, 2005). The way a timetable is constructed can directly affect the level of service 46 for passengers, both in terms of the frequency of the services offered as well as the ability 47 for trains to run on time. In dense areas, trains operate with minimal headway, based on 48 the block lengths, characteristics of the train such as running and braking speeds, and local 49 guidelines (Goverde, 2005; Palmqvist, 2019). Some margin has also to be applied. When 50 tolerances are planned too tight, i.e. too little time is scheduled to realistically perform 51 the tasks at hand, the likelihood of delays increases. On the other hand, if too much time 52 is planned for each task, the frequency at which trains can operate will be reduced and 53 with it the attractiveness of railways. Timetabling thus poses a balancing act in which is 54 it important to schedule both realistic and appropriate times for all the different elements 55 that make up the final timetable (Hansen, 2010), a non-trivial task. Broadly speaking, a 56 timetable for passenger trains can be divided into two main elements, these being the run 57 and dwell time of a train. The run time refers to the time it takes for a train to travel 58 between two subsequent stations, whereas dwell time refers to the time a train is stationary 59 at a station. The study presented here focuses on the latter of these, the dwell time. 60

61 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 62 dense railway networks (about 20% in the Paris region, for example), where trains halt 63 often to allow for passengers to board and alight. Furthermore, dwell times can have a 64 strong effect on the operation of trains in dense areas. Since these trains are bound by 65 this minimum headway the dwell time can become the limiting factor to determine the 66 frequency at which trains can be operated. When dwell times are not accurately scheduled 67 this can lead to dwell time delays. Dwell time delays arise when a train is stationary 68 for longer than scheduled (Buchmueller et al., 2008) and although small can accumulate 69 over an entire journey (Christoforou et al., 2020) and cause knock-on delays due to trains 70 occupying a platform for longer than scheduled (Yamamura et al., 2012). Some studies 71 proposed estimation of dwell times via short-term prediction (Li et al., 2016) but only for 72 short stops with no mandatory departure time. A major challenge for railway operators 73 in high-density networks is to define adequate dwell times when designing a timetable: a 74 dwell time which both ensures a feasible timetable (Hansen, 2010) and an optimal use of 75 the available capacity (Goverde, 2005). 76

77 1.3 Scheduling of dwell times

Despite the importance of dwell times, current approaches to scheduling dwell times rely 78 on rules of thumb and the experience from planners (Wiggenraad, 2001; Christoforou et 79 al., 2020). Scheduling dwell times will become increasingly important as operators want 80 to increase the frequency of trains operated, often without increasing the available infras-81 tructure. In the Netherlands, for example, there is the ambition to run six trains an hour 82 on some busy corridors (IenW, 2022), putting increased stress on the need for adequate 83 scheduling principles. An important aspect of scheduling dwell times is to know the dwell 84 time without margin referred to as tight dwell time in this work. The tight dwell time is 85 especially relevant for timetabling in heavily utilized railway systems, as the dwell time 86 becomes a limiting factor (see for example the impact of the crowd in simulations done by 87 Luan & Corman (2022)). 88

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

Other authors have tried to approach this notion, referring to various heuristics as minimum dwell times which correspond to the minimum time needed to complete the alighting and boarding process and to depart from a station if everything goes as fast as possible (Coverde 2005: Pederson et al. 2018). Although several dwell time models that attempt to

(Goverde, 2005; Pedersen et al., 2018). Although several dwell time models that attempt to

predict such a minimum dwell time exist, mainly based on regression models, the potential
of such models has not yet been realized (Yang et al., 2019). For example, testing the
well-cited dwell time model by Weston (1989) at a larger number of stations, Harris (2005)
found that the model does not perform well under high passenger loads.

Furthermore, many of the studies focusing on minimum dwell times are limited by small 99 sample sizes, limited to a small number of stations, or limited to manual observations 100 (see for example the work done by Puong (2000)). Recent years have seen an increase in 101 the volume and granularity of the available data (Palmqvist, 2019) allowing for a more 102 in-depth study into dwell times. At the same time, research into minimum dwell times 103 has remained scarce, with a recent example being the study by Cornet et al. (2019) who 104 made use of automatic passenger count data to infer a minimum dwell time. Moreover, 105 the definition of minimum dwell time indicates that this is the dwell time corresponding 106 to when everything goes as fast as possible. Such an assumption is not realistic, that is 107 why we introduce the notion of tight dwell which instead is an estimate of the necessary 108 dwell time for people to alight and board properly without any extra time. To the best of 109 our knowledge, no studies have dealt with a similar notion so far. 110

111 1.4 Objective of the paper

Given the need for an increased understanding of tight dwell times, the objective of the 112 study presented here is to define a way in which the tight dwell time can be better estimated 113 for commuter trains. To do so we make use of highly detailed passenger flow data collected 114 on board suburban commuter trains in the Paris region area and propose two novel methods 115 to estimate tight dwell times based high resolution passenger flow data. The first method 116 makes use of the time gaps between the flows of passengers, defining the tight dwell time 117 based on clusters of passengers. The second method makes use of the cumulative flow of 118 passengers. We then define which of the proposed approaches is most promising, compare 119 it to methods found in the literature, and show the usability and precision of our proposed 120 methods. It is worth noting that we do not attempt to define or estimate the dwell time 121 margin itself, as this requires information on the scheduling principles which cannot be 122 disclosed. Instead, the work presented here is limited to the tight dwell time itself. 123

124 1.5 Structure of the paper

The remainder of the paper is organized as follows. First, we present a literature review 125 (Section 2), in which we define the components of dwell times precisely and how dwell time 126 has been estimated up to now. Some formalization of the problem components follows 127 (Section 3). Our data is then described (Section 4). Our two new methods to estimate 128 the tight dwell times are presented in Section 5, along with two existing methods used to 129 estimate minimal dwell times. We thus compare their results (Section 6). This comparison 130 leads us to select the method based on cumulative flows, and by comparing it further with 131 the two selected methods from the literature, we show its superiority when sufficiently 132 high-resolution data is available. Section 7 illustrates the most appropriate method on 133 a branch of line N of the Paris suburban network and shows that we can even make a 134 door-by-door analysis. A discussion and a conclusion end the paper. 135

136 2 LITERATURE REVIEW

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



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

140 2.1 Components of dwell times

Although the dwell time of a commuter train is often spoken of as a single process it is, 141 in fact, made up of different individual processes. Seriani & Fujiyama (2019) define dwell 142 time as having both static and dynamic time elements. According to their definition, the 143 static element of a dwell time consists of mechanical processes such as the door opening 144 and closing, and the dynamic element consists of the alighting and boarding process. In 145 line with this, Buchmueller et al. (2008) state two main processes make up the dwell time, 146 these being the passenger service time and the train dispatching time. The authors further 147 divide dwell time into five different sub-processes, including the door-unblocking, door 148 opening, boarding/alighting, door closing, and train dispatching process. 149

The time needed to complete the static elements of the dwelling process is dependent on 150 the design of the train. Trains can be equipped with sliding extensions to allow for a level 151 entry, for example. The time needed to open the door is longer when such an extension is in 152 place (Buchmueller et al., 2008). Although considered to be a static element of dwell times, 153 the time needed to open a door can be dynamic when the doors do not open automatically 154 and passengers are required to request a door opening (Harris, 2015). The time needed to 155 complete the dynamic element of dwell times is governed by the alighting and boarding 156 time (Goverde, 2005). Since trains have multiple doors there are multiple alighting and 157 boarding processes at the same time, and it is the door where this takes the longest which 158 defines the alighting and boarding time making it the "critical door" (Buchmueller et al., 159 2008; Coulaud et al., 2023). Note that the critical door is specific to the station, the time 160 of the day, the rolling stock, and so on. This adds another source of variability, which 161 concurs in justifying the necessity for margins. 162

163 2.2 Margins and robustness

When designing a timetable it is important to include realistic running and dwell times to 164 ensure that the final timetable is feasible (Hansen, 2010). One way to ensure the feasibility 165 of the scheduled times is to include some form of margins in both the running and dwell 166 times. Here, margins refer to time supplements added on top of the scheduled times to 167 allow for any fluctuations, be it a result of driver behavior or delays (Andersson et al., 168 2011; Goverde, 2005). These margins are important since it is unlikely that the actual 169 operation will follow the scheduled times. Adding margins to running times is common 170 in timetables (Palmqvist, 2019), and are usually a percentage of the nominal running 171 time. This percentage is based on local planning principles, with seven percent being the 172 norm in the Netherlands (Goverde, 2005), three percent in Sweden (Palmqvist, 2019), and 173 five percent for the suburban trains operated in Paris. Other margins can be added at 174 important locations, also known as nodes, and is common practice in Sweden (Palmqvist, 175

¹⁷⁶ 2019) and Switzerland (Vromans, 2005), for example, and is also common practice in ¹⁷⁷ France.

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

185 2.3 Estimating tight dwell times

As mentioned, margins can be added to both the running and dwell time of trains. Running 186 time margins are relatively straightforward to calculate since the nominal running time can 187 be calculated based on a feasible speed-distance profile over an open stretch of track for 188 a train, taking the track alignment characteristics into account (Goverde, 2005). The 189 nominal running time can then be compared to the realized operation of a given set of 190 trains to measure the margins present in the timetable. In contrast to this, dwell time 191 margins are less straightforward to calculate given the stochastic nature of dwell times. 192 The alighting and boarding time, especially, makes the actual dwell time subject to high 193 variability (Cornet et al., 2019), and comparing the scheduled and realized dwell times is 194 not trivial. 195

Only a limited number of methods have been proposed in the past to measure the margin 196 present in dwell times. One way to do so is to make use of situations where a train 197 arrives with such a delay that it enters the station after its scheduled departure time. An 198 example of this approach was explained by Pedersen et al. (2018). Coulaud et al. (2023) 199 also highlighted that late train dwell times are impacted at the first order by the passenger 200 flows which suggests a reduction of margin for late trains. In such cases, trains will depart 201 as soon as possible, and the dwell time is likely to only include the alighting and boarding 202 time. Yet, none of these approaches quantified the reduction in margins for late trains. 203

Some authors (Cornet et al., 2019) criticized this approach highlighting that some late 204 trains could not leave the station right after the alighting and boarding process was com-205 pleted, due to signalization and dispatching decisions for instance. Daamen et al. (2009) 206 and Keeman & Goverde (2013) proposed methods to identify route conflicts that could 207 be used for filtering out trains without conflict. For example, Kecman & Goverde (2015) 208 presented a machine-learning approach to predict dwell times using this filtering method. 209 Apart from route conflicts, the driver's behavior and his sensitivity to how much time 210 to wait before leaving may impact such estimation of the alighting and boarding time. 211 Another approach to overcome these problems is proposed by Cornet et al. (2019) taking 212 advantage of passenger flow data instead. Having access to detailed passenger count data 213 (i. e. the total number of people alighting and boarding the train and the load of the 214 train at every stop) Cornet et al. (2019) defined a measure they call the reduced passen-215 ger flow. The minimum dwell time is subsequently obtained as a function of the reduced 216 passenger flow. Whilst promising, the two approaches described above are only defined for 217 late trains and do not allow to capture the variability in the flow of both alighting and 218 boarding passengers. Moreover, only the study of Coulaud et al. (2023) used door-by-door 219 counting data but focused on dwell time prediction. In this work, we are presenting a way 220 to estimate tight dwell times for any stop, a metric that captures the variability of the 221 passenger flow. 222

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

| 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 | $\sum_{t}^{j} N_{k,s,d,j,t}$ |
| Alighting or boarding at time t | Number |
| (relative to the latest opening of the door) | 1 * <i>k</i> , <i>s</i> , <i>d</i> , <i>J</i> , <i>t</i> |

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

223 **3** FORMALIZATION OF THE PROBLEM COMPONENTS

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

226 3.1 Hierarchy of resolutions in passengers flow data

In this section, we introduce the different levels of resolution (and the related indexes) that 227 may exist concerning the available passenger flow data. The highest resolution for counting 228 passengers that either alight or board a train is the movement of passengers through a door. 229 A passenger is defined as someone who alights or boards a train once. In practice, this 230 means that a passenger who has boarded a train and then alights the train later will be 231 counted twice, once as a boarding and once as an alighting passenger. Count data on this 232 resolution does, however, not provide information on when and where the event took place. 233 To be able to do so, four pieces of information are needed: the train number k, the station 234 s, the day d, and the door j. An additional fifth piece of information is required to be able 235 to study dwell time components, this being the time when a passenger crosses the door 236 relative to the opening of the door t. Table 1 sums up these notations. 237

The highest resolution of data can be noted as $N_{k,s,d,j,t}$, and provides information on 238 the number of passenger movements for train k, at station s, on date d, at door j, and 239 time t after the opening of the door. Lowering this resolution is often synonymous with 240 aggregating $N_{k,s,d,j,t}$ over several indexes that correspond to operational quantities. Table 241 2 shows several examples of levels of resolution. For example, summing $N_{k,s,d,j,t}$ over time 242 (t) and door (j) results in the volume of passengers per stop; and summing $N_{k,s,d,j,t}$ over 243 time (t), door (i), and train number (k) results in the daily station volume. Summing over 244 all the indexes means getting the total volume of passengers for the studied line in the 245 studied period. 246

Although higher resolutions enable more precise analysis, it is harder to reach in practice.
Indeed, three kinds of restrictions may exist to access higher resolutions of passenger flow

249 data:

1. Data collection. If the data is collected manually, it is not possible to be precise enough to get the time t of the passengers' movement relative to the opening of the 252 door;

253 2. Data storage. If the data is collected automatically, the data has to be transmit-254 ted from the train to a server and stored. The highest resolution makes the data 255 heavier and the storage on-board and transmission off-board are either complicated 256 or expensive which is dissuasive for many transport companies;

Competition between companies. Knowing precisely the passengers' movements
 is a critical piece of data for railway operation and, provided they have higher resolution data available, some companies may consider it confidential.

As a consequence, most of the works done using counting data so far have used lower-260 resolution data. Figure 2 illustrates the data hierarchy by showing the levels of resolution 261 and associated use cases and works. At a macroscopic scale, the Global Transport In-262 vestigation (DRIEA, 2022) shows, among other information, the evolution of the total 263 volume of passengers in the Paris region every decade. The annual station volume is also 264 interesting to model, especially for new stations or new lines, the four-step model is the 265 most common tool to do so (McNally, 2007). The passenger volume of one day can be 266 a useful piece of data to predict disruption for example (Yap & Cats, 2021). When only 267 manual counting data is available, the annual volume for a train and a station is the most 268 appropriate piece of data to build timetables (Brethomé, 2018). Besides, the daily station 269 prediction (De Nailly, 2023) does not need more than daily station volumes. The vol-270 umes by stop helped improve the robustness of timetables (Cornet et al., 2019; Medeossi 271 & Nash, 2020; Christoforou et al., 2020; Kuipers & Palmqvist, 2022) while the volume by 272 door is a precious piece of data to model dwell times (Coulaud et al., 2023; Kuipers, 2024). 273 The highest level of resolution has not been tackled in the literature, to the best of our 274 knowledge. 275

276 3.2 Dwell time components

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

$$DT = ABT + M + TT.$$

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

284

2

$$TDT = ABT + TT$$

We can suppose the technical time (TT) to be constant among the doors and the stops 285 (Buchmueller et al. (2008) showed some variability but was less important than the one 286 observed for the other variables). By contrast, the alighting and boarding time can vary 287 among the stops and the doors for a given stop and the margin can differ between stops as 288 well. The alighting and boarding time is the necessary time for all the passengers to either 289 alight or board the train. As an example, the time for a passenger arriving late at the 290 station and boarding right before the departure of the train is included in the alighting and 291 boarding time while a long period between two passengers' movements is not included. We 292 define the alighting and boarding time for a given door i as a door alighting and boarding 293 time $(DABT_i)$, and the margin as a door margin (DM_i) . We then have: 294

$$ABT = \max_{j} DABT_{j}$$

296
297
$$M = \min_{j} DM_{j}$$



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.

| Variable | Notation |
|--|-----------|
| Dwell Time | DT |
| Alighting and Boarding Time | ABT |
| Margin on the observed dwell time | M |
| Technical Time | TT |
| Local Alighting and Boarding Time (door level) | DABT |
| Local Margin (door level) | DM |
| 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.

298 3.3 Cornet's minimum dwell time definition

Cornet et al. (2019) introduced the notion of minimum dwell time (MDT) for a reduced passenger flow (p), for a fixed station s_0 . Here, p represents the size of the passenger flow and has a value between 0 and 1 such that p = 0 corresponds to a train with almost no passenger exchange and p = 1 to a train with many alighting and/or boarding passengers. The minimum dwell time for a fixed reduced flow p_0 is obtained as follows:

304
$$MDT(p_0, s_0) = \min_{k, d \ s.t. \ p=p_0} DT_{k, s_0, d}$$

In the remainder of this paper, we estimate MDT and TDT with different methods, the notation \widehat{X} is used for the estimation of any variable X. We will then compare the results of our new methods with the ones of Cornet's method. Yet, we do not estimate the same variables. However, by adding the hypotheses that \widehat{ABT} and \widehat{M} are independent and that $\widehat{ABT} = \lim_{k,d} |\widehat{A}_{k,d}| \to \infty \left(\min_{k,d \ s.t. \ p=p_0} \widehat{M}_{k,s_0,d} \right) = 0$, we can demonstrate that:

310

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

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

This proposition (demonstrated in Appendix) makes it possible to perform a comparison between \widehat{MDT} as defined by Cornet et al. (2019) and \widehat{TDT} that we estimate based on the methods presented in this study.

320 4 PRESENTATION OF OUR DATA AND PERIMETER

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

323 4.1 High resolution passenger flow data used in this study

The rolling stocks in operation are equipped with an automatic passenger count system at each door which counts both the number of alighting and boarding passengers using infrared sensors. Every several seconds, a signal is sent to an onboard computer that registers the cumulative number of passenger movements. In addition to this, the system



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



Figure 4: Regio2N design for line N of the Paris commuter trains network (top). Data at the counting event scale 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.

registers whether the train doors are open or closed. It also provides information on the train number and the name of the stop. An important number of stops (around 30%, randomly distributed) is not counted by the sensors which leads to gaps in the data. Such gaps are not issues for our study since we are only focusing on the stops where the data is available.

The data used in this study originates from *Line N* of the Paris commuter trains network, 333 shown in Figure 3 and has been collected over six months from September 2022 to March 334 2023. Line N consists of three branches and various services are operated on each branch 335 (omnibus and semi-directs). The rolling stock in use is the Regio2N, illustrated in Figure 336 4. Each carriage has eight doors which are 1.60m wide and require passengers to press 337 a button to initiate the opening procedure. The technical time is supposed deterministic 338 and uniform at 7s (as reported in the technical documentation). The services on this line 339 are operated with both single units and double units. 340

The study presented here (see Section 6) focuses on Versailles-Chantiers, which is the main 341 station in line N (except for Paris terminus) both in terms of passenger traffic (nearly 15 342 000 passengers boarding per day in 2018) and in terms of train traffic (more than 200 343 trains a day on weekdays). To compare the methods, we select only trains toward Paris. 344 A more general analysis is performed in Section 7 where we consider all the services on 345 Dreux branch. On this branch, two trains per hour during peak hours and one train 346 per hour during off-peak hours are operated. This branch undergoes two convergences 347 with the other branches of the line: the first between Villiers-Neauphle-Pontchartrain and 348 Plaisir-Grignon, the second between Plaisir-Grignon and Versailles-Chantiers. 349

The automatic passenger count data makes it possible to have access to data close to the 350 highest resolution we presented previously. The data is disaggregated in terms of k, s, d, j351 and partially aggregated on t (as shown in Table 1). Indeed, for each quadruplet (k, s, d, j), 352 we have a list of counting events. A counting event means that at a given time t', $N_{k.s.d.j.t'}$ 353 passengers have alighted/boarded the train since the latest counting event (or the opening 354 of the door in the case of the first counting event). Figure 4 shows an example of such 355 kind of data. The time separating two counting events is truly random, having a standard 356 deviation of 18s for an average of 13s and a median of 7s. 357

The daily alighting and boarding volumes for a given train number are compared to determine the accuracy of the system. For a given train k at a given date d, the total number over all stops of people alighting shall be the same as the total number of people boarding. In the case of the presented data, the total number of alighting differs by 5% on average (3% in median) from the total number of boarding which gives an insight into a decent data quality. In addition to this, a study by the provider of the data has validated the quality of these sensors.

365 4.2 Dwell time measurement

For dwell time measurement, several kinds of sensors measure the observed dwell times via 366 automatic vehicle localization (AVL). In this study, we used data from the infrared sensors. 367 These sensors indicate the time when the door is fully open and the time when the door is 368 fully closed. This data has been validated by the operator. Traditionally, dwell times are 369 measured as the time between the last rotation of the wheels when the train arrives and 370 the first rotation of the wheels when the train leaves. However, the first rotation of the 371 wheels may happen quite a long time after all doors are closed (about 6s) and may vary 372 from one train to another and from one driver to another. As the opening and closing of 373 the door take deterministic times, our choice helps reduce the randomness included in the 374 technical time when comparing a method estimating the alighting and boarding time and 375 a method that measures a dwell time (see Section 5). Moreover, it helps avoid potential 376 synchronization issues between clocks of different sensors. 377

378 5 Presentation of methods to estimating tight dwell times

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

383 5.1 A simple heuristic: observed dwell times for late trains

One way to determine the tight dwell time is to make use of situations where the train 384 enters the station with such a delay that it arrives after its scheduled departure time. An 385 example of this approach was explained by Pedersen et al. (2018). They hypothesized that 386 a late train would only dwell for the shortest time needed, i.e. the time needed for the 387 alighting and boarding process to be completed. This is commonly done to recover some 388 of the delay time. The margin is then at least partly consumed as shown by Coulaud et al. 389 (2023). In the sequel, we will compare the dwell time in this situation to the tight dwell 390 time that we will estimate with our new method. 391

One particularity of this method is that it does not need any passenger flow data which 392 makes it an attractive method: it only relies on the scheduled and actual arrival and 393 departure times. Nevertheless, it does come with a drawback; Cornet et al. (2019), for 394 example, criticized this approach highlighting that some delayed trains could not leave the 395 station right after the alighting and boarding process was complete due to signalization 396 and dispatching decisions for instance or relative to the driver behavior. This could lead 397 to situations where dwell time is significantly greater than the tight dwell time. Some 398 works proposed corrections to improve this approach, filtering route conflict (Daamen et 399 al., 2009; Kecman & Goverde, 2013, 2015) but for the sake of simplicity, we will not further 400 discuss these route conflict corrections in the sequel. 401

402 5.2 A heuristic based on passenger volumes: Cornet's minimum dwell time

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

- 406 1. Select one station and one direction;
- 407 2. Perform a principal component analysis (PCA) on the number of alighting at one 408 stop, the number of boarding at one stop, and the load of the train to define p (the 409 normalized first principal component) called reduced passenger flow;
- 3. Consider only trains for which the observed departure time is strictly greater than
 the scheduled departure time;
- 4. Separate all the stops by their values p in 200 windows ([0,0.005];[0.005,0.01]...);

5. Get the minimum dwell time \widehat{MDT} as the minimal observed dwell time among the window of p in which the stop is.

In this case, p is an illustration of the relative importance of the passenger flow and is a value between 0 and 1 with p = 0 meaning that there are no passengers, p = 1 indicating the maximum flow encountered in the dataset.

In link with the available resolution of data $(\sum_{j,t} N_{k,s,d,j,t})$, Cornet et al. (2019) assumed that the flow was uniformly distributed among the doors, something which is not likely. Given that the data available for our study provides information on the distribution of passengers on a door-by-door level, it is possible to include this aspect. Another adaptation to the method proposed by Cornet et al. (2019) is the exclusion of the load factor of the



Figure 5: 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.

train. The line studied here is not heavily used and it is possible to assume that standingpassengers are rarely present given the seating capacity of the trains.

These adjustments mean that we changed step 2 and step 4 of the workflow presented above. In step 2, we performed a PCA on the number of people alighting at one door and the number of people boarding at the same door. In step 4, we consider for each stop the maximum value p among the doors before separating the stops. Figure 5 shows the last step of this framework for *Versailles-Chantiers* toward Paris.

This method is based on the hypothesis that *MDT* entirely depends on the reduced passenger flow *p*. However, one may conjecture the existence of instabilities in the alighting and boarding time regarding the passenger flow such as slower passengers and passenger boarding before the end of the alighting process.

434 5.3 Tight dwell time estimation using high-resolution passenger flow data

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

442 5.3.1 Cluster method

The cluster method is inspired by an existing method for the estimation of the alighting 443 and boarding time at the door level (Wiggenraad, 2001). The main challenge to estimate 444 DABT is to distinguish the main flow of passengers from the margins. Indeed, late arriving 445 passengers or people remaining on the platform waiting for the alarm to board may bias 446 our understanding of the alighting and boarding process. Then, one needs to make the 447 difference between isolated movements and cluster movements. For Wiggenraad (2001), a 448 cluster is the movement of a group of passengers separated by less than three seconds (the 449 cluster size parameter, Δ_t). By contrast, an isolated movement would be the movement of 450 one person separated by more than three seconds from the other passengers. 451

452 Such a protocol could be put into practice without modification if the data had the highest



Figure 6: 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.

resolution. In our case, we deal with counting events, so we need to adjust our method. Thus, we take the following definition: two counting events are in the same cluster if there are less than Δ_t seconds by passengers in the later counting event relative to the time between the two counting events. We define the estimated $D\widehat{ABT}$ as the duration of the first cluster. Figure 6 illustrates the process of estimating $D\widehat{ABT}$ with the cluster method. As explained in Section 3, to get \widehat{ABT} one needs to take the maximum $D\widehat{ABT}$ among the doors of the train for this stop $(\widehat{ABT} = \max_j D\widehat{ABT}_j)$. The estimated tight dwell time

460 will therefore be $\widehat{TDT} = \widehat{ABT} + TT$.

The value of the cluster size parameter (Δ_t) shall be chosen carefully. Indeed, a too-461 low value for Δ_t would lead to splitting clusters in isolated movement, resulting in an 462 underestimation of \widehat{TDT} . A too-high value for Δ_t would imply considering too many 463 passengers in clusters and part of the margin would be included in the estimated TDT. 464 As an illustration, Figure 7 shows the mean value of \widehat{TDT} for various Δ_t values (among all 465 the trains toward Paris at Versailles-Chantiers). This confirms our qualitative analysis as 466 TDT increases with Δ_t until reaching a plateau. In the sequel, we are focusing on $\Delta_t=3s$ 467 and $\Delta_t = 7s$ (respectively before and after the plateau). 468

469 5.3.2 Quantile method

Another way to deal with isolated movements is to consider the necessary time for a 470 quantile q of the passenger flow to perform the process. This necessary time has, then, 471 to be expanded to take into account all the passengers proportionally to the portion q of 472 them. Figure 8 illustrates this method for q=0.9. On the left of the figure, we can see 473 that the last passengers' movements are realized much later than the other passengers' 474 movements. Then, we extrapolate the necessary time for this quantile q (orange cross) to 475 obtain the time per passenger to board/alight the train (i.e. the slope of the green line). 476 We then take DABT (purple line) as the product of this time per passenger by the total 477 number of passengers which corresponds to the intersection between the green line and the 478 red line. The idea behind this method is to give every passenger a sufficient amount of time 479 to alight/board the train without adding too much time for the slowest passengers. Once 480 again, this method defines an estimation of $D\widehat{ABT}$, and one needs to take the maximum 481 among the doors and add the technical time to obtain the desired estimation of TDT. 482

The choice of q is important. A too-low value would make the extrapolation less relevant as the average time per passenger is taken on fewer passengers. A too-high value leads



Figure 7: Average estimated tight dwell time (\widehat{TDT}) over the dataset computed with the cluster method versus the cluster size parameter Δ_t . The higher Δ_t the higher \widehat{TDT} on average until reaching a plateau.



Figure 8: Cumulative passengers (alighting and boarding) against time (blue dots) and quantile q=0.9 (the orange cross), (left) $D\widehat{ABT}$ 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 $D\widehat{ABT}$ corresponding to the intersection of both lines is in purple (right).



Figure 9: Average estimated tight dwell time (TDT) over the dataset computed with the quantile method versus the quantile q. High values are obtained for little q due to side effects. After reaching a minimum, the value increases again with q corresponding to the fact that slower passengers alight/board later in the process.

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

Table 4: Subsets required to use the different methods.

to considering several people who are not in the main flow in the computation. Figure 485 9 shows the mean value of TDT for various q values (among all the trains toward Paris 486 at Versailles-Chantiers). One may see that the highest \widehat{TDT} on average are obtained 487 for q=0.1 corresponding to the reaction time of the first passengers at the very beginning 488 who take longer to alight/board. By contrast, the lowest TDT are obtained on average 489 for q=0.5 which is coherent with the fact that people moving first (alighting or boarding) 490 tend to be faster than the others. In addition, most of the time the first passengers 491 movements correspond to alighting movements which are often faster than boarding ones. 492 After q=0.5, the average TDT increases until q=0.9. In the sequel, we are focusing on 493 q=0.4 (corresponding to taking less than half of the passengers for the extrapolation) 494 and q=0.9 (close to all the passengers) and comparing this method to the three already 495 presented. 496

497 6 Results: Comparison of the methods

In this section, we show the comparisons between the four methods presented previously. In the following subsections, the most restrictive sample is chosen to compare the methods two by two where both are defined. Table 4 reminds the sub-sets on which each method is defined and highlights the subset used in the following subsections.

502 6.1 Quantile method vs Cluster method

First, we compare the quantile method and the cluster method. Figure 10 shows the 503 comparison between these two methods with the sets of parameters $(\Delta_t=3s, \Delta_t=7s)$ and 504 (q=0.4, q=0.9). We see a good correspondence between the two methods most of the time. 505 However, when there is a significant difference between the tight dwell times obtained by 506 the two methods it is often a case of underestimation of the cluster method compared to 507 the quantile method. Indeed, the cluster method ignores the isolated passenger movements 508 while everyone is provided a time to alight/board in the quantile method. For example, if 509 the flow is slow enough: the cluster method may ignore most of the passengers while the 510 quantile method may add a particularly long time to the duration of the quantile. This 511 results in a large difference between the two estimations. The case of $\Delta_t=7s$ and q=0.4512 is specific, we see a light underestimation of the quantile method in many cases. Taking 513 q=0.4 gives on average close to the smallest TDT (see Figure 9 in Section 5.3.2) while 514 $\Delta_t = 7s$ gives on average values close to the highest (see Figure 7 in Section 5.3.1). In 515 other words, $\Delta_t = 7s$ means ignoring very few isolated passengers' movements and q = 0.4516 computes the time per passenger where it is the smallest. 517

Without excluding the cluster method from practical estimations, the quantile method is more resilient in terms of data quality. Indeed, the times of the counting events are only used in a cumulative way which mitigates the impact of potential gaps in the data which is not the case for the cluster method. Some points, Figure 10, illustrate this fact, they have particularly low \widehat{TDT} with the cluster method while having quite a high \widehat{TDT} with the quantile method. For the sake of consciousness, we will compare the existing methods only with the quantile method in the following sections.

525 6.2 Estimated tight dwell times with quantile method vs observed dwell times 526 for late trains

In this section, TDT estimated with the quantile method (presented in Section 5.3.2) are 527 compared to observed dwell times (DT) for late trains (presented in Section 5.1). This 528 comparison, shown in Figure 11 is performed using one station (Versailles-Chantiers) and 529 one direction (toward Paris), and only trains for which the observed arrival time is greater 530 than the theoretical departure time are selected. Observed dwell times for late trains are 531 generally greater than TDT no matter their value. This means that train drivers are 532 not able to leave the station as soon as the alighting and boarding process is completed. 533 DT for late trains is often much greater than TDT computed with the quantile method 534 and the difference is less for q=0.9 than for q=0.4. This difference is coherent with the 535 fact that slower passengers' movements are taken into account with a greater value of 536 q. Nevertheless, several points do not follow this general observation and their computed 537 $\overline{T}D\overline{T}$ with the quantile method is greater than their actual DT. This is an artifact of the 538 quantile method: if the chosen quantile q includes a high proportion of slower passengers, 539 then TDT will be over-estimated with the quantile method. This situation is rare and 540 even rarer with a higher q value as slower passengers need to be more numerous at the 541 given stop in those cases. 542

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

In this section, 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 theoretical departure time, which makes Cornet's method applicable. Figure 12 shows the histograms of \widehat{TDT} values for two values of p and two values of q 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



Figure 10: Comparison of estimated tight dwell times (TDT) between the quantile method with q=0.4 (top) and q=0.9 (bottom) and cluster method with $\Delta_t=3$ (left) and $\Delta_t=7$ (right). Both methods give most of the time similar 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.



Figure 11: 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=0.4 (left) and q=0.9 (right). Except for rare exceptions, DT is almost always larger than \widehat{TDT} for late trains.



Figure 12: 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=0.4 (left) and q=0.9 (right). Except for p=0.05 and q=0.9, 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.

method, except for the case of p=0.05 and q=0.9 where the minimum is greater but close 551 to Cornet's MDT. We can interpret this result as suggesting the existence of residual 552 margins in Cornet's estimation or an underestimation in the quantile method. The spread 553 of TDT values is non-negligible for one value of p while MDT is unique which means that 554 the quantile method can take into account the variability of the flow. 555 Figure 13 shows the MDT (estimated with Cornet's method) and minimal TDT (obtained 556 with the quantile method) versus the reduced flow p. We cannot estimate \widehat{MDT} for high p. 557 The values obtained with the quantile method are always equal or less than Cornet's MDT. 558 As already mentioned, nothing can guarantee Cornet's MDT to be free of margin while the 559 tight dwell time obtained with the quantile method is defined as free of margin. Moreover, 560 a quantile of the flow is not the total flow: for instance, if people in the quantile are much 561 faster than the others, this will lead to an underestimation of TDT (and vice versa). Hence, 562 the minimal value of \widehat{TDT} for a given p can be obtained for a particularly fast quantile 563 compared to the rest of the flow which could concur in explaining the difference. 564



Figure 13: 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=0.4 (left) and q=0.9 (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.

565 6.4 Summary of the comparisons

In this subsection, we summarize the previous results. We review the distribution of the value obtained with each method, fixing the most appropriate parameter for the cluster and quantile method is required to do so. To make everything comparable, we consider only late trains (those having their observed arrival time greater than their theoretical departure time). After that, we assess the methods based on new qualitative indicators. These indicators evaluate how relevant each method is for practical use.

As seen in Section 5, the quantile method is quite robust to the parameter choice from 572 q=0.2: despite small differences, the average of the estimated tight dwell time is close. 573 Qualitatively, a better estimation is performed when the time per passenger is computed 574 on more passengers (which means a higher q) but a too-high value may take into account 575 particularly slow passengers. We can use an elbow criterion in Figure 9 (in Section 5.3.2) to 576 decide which value to choose. In our case, the most appropriate value based on this criterion 577 would be q=0.8. A similar comment can be stated for the cluster method which is also 578 robust to the parameter choice from $\Delta_t = 2s$, having similar average values. Qualitatively, 579 a too-low value would result in ignoring too many passengers whereas a too-high value 580 would result in considering the entire flow as the main flow. We can, once again, use an 581 elbow criterion in Figure 7 (in Section 5.3.1) to decide which value to choose. In our case, 582 the most appropriate value based on this criterion would be $\Delta_t = 6s$. Those two values are 583 used for a summary of the comparisons between all the methods discussed so far. 584

First, Figure 14 shows the statistical distribution of the values obtained with each method. We see that for late trains, all the methods give values smaller than the actual observed dwell time DT. Cornet's \widehat{MDT} is close to the values of both the cluster and the quantile method but with a less important range between small and high values. The quantile method gives \widehat{TDT} slightly greater than the cluster method.

Nevertheless, this paper aims to show how relevant those methods are for practical use.Table 5 qualitatively sums up the pros and cons of each method over the following criterion:

- The robustness relative to data lack of quality,
- The resolution of the data used (see Section 3.1),
- The range of usability if the data is accessible,
- The amount of necessary data to perform one estimation.



Figure 14: 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.

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

The quantile method has high robustness and resolution; it has a large range of usability when the data is accessible. The amount of necessary data is the only item where the method is challenging. Nevertheless, the data required is provided continuously by the APC system which makes it easy to access once the system is available. To conclude, the quantile method is the most appropriate way to estimate a tight dwell time for practical implications.

$_{602}$ 7 Use case on the line N

In this section, we apply the quantile method to estimate the tight dwell times for all 603 stations of the branch Paris-Dreux on line N (presented in Section 4.1). We chose to focus 604 only on the direction toward Paris and on weekdays. The terminus (Paris-Montparnasse 605 and *Dreux*) were removed since there is no point in estimating a tight dwell time at a 606 terminus. We selected the couples (k,s) where we had at least ten dates of observations 607 (d) and could then estimate a tight dwell time. This resulted in the absence of certain 608 train numbers in our analysis due to gaps in the data (see Section 4.1). Figure 15 shows 609 the average values of TDT for all the stations and all the services of this branch and this 610 direction computed with the quantile method (q=0.8) between September 2022 and March 611 2023. The stations where \widehat{TDT} was on average higher correspond to the main stations of 612 the line (i.e. Versailles-Chantiers and Plaisir - Grignon). These two stations also being 613 the only connection stations in the line (except for the terminus), some services also had 614 higher TDT on average. These services corresponded to the morning peak hours. Line N 615 being a commuter train, the direction toward Paris was the busiest during the morning 616 peak hours, which was coherent with having higher tight dwell times. One can notice 617 that the station Houdan had quite high TDT on average, whereas it was not a connection 618 station nor did it gather some specific point of interest. So far, we have no explanation for 619 these values. 620

Figure 16 shows the same illustration for the opposite direction (toward Dreux). The 621 observations on the connection stations (Versailles-Chantiers and Plaisir - Grignon) still 622 hold. The busiest services were in the evening peak hours which is coherent with line 623 N being a commuter train. However, the mean values of TDT estimated in the main 624 station were smaller toward Dreux. This is possibly due to a lack of train numbers in our 625 data, especially in the evening peak hours. If the busiest trains are not recorded in the 626 evening peak hours, it is normal to see a difference with the morning peak hours. We can 627 also imagine that the distribution of the passengers' movements might have been more 628 homogeneous in the evening than in the morning. 629

This second assumption is partly validated by Figure 17 which shows the mean door 630 alighting and boarding time DABT estimated for the Paris-Dreux services at Versailles-631 *Chantiers.* We find that the front doors are often the busiest toward Paris with some doors 632 with a particularly long DABT (door 2 for service 164866 or doors 7 and 8 for service 633 164878). Toward Dreux, the mean DABT values are more homogeneous. For services 634 165847 and 165855 in particular, the maximal DABT is not that high on average while 635 being important for several doors. Toward Paris, doors 1 to 8 are always busier than doors 636 9 to 16, this is probably because the terminus *Paris-Montparnasse* is a dead-end station 637 where the exit is at the front of the train. The same explanation may hold for trains toward 638 Dreux where the pattern is the opposite (doors 9 to 16 are busier because they are closer 639 to the entrance of the platform). 640

A widespread idea to explain the length of dwell times is that alighting and boarding times are linear with passengers' movement. Figure 18 mitigates this general idea, showing the estimated $D\widehat{ABT}$ with the quantile method (q=0.8) versus the number of passenger movements at Versailles-Chantiers, toward Paris, at door 8. Despite an increasing trend, we see a large spread of the $D\widehat{ABT}$ for a given number of passenger movements. This is



Figure 15: Average estimated tight dwell times (\widehat{TDT}) for all services on line N, branch Paris-Dreux toward Paris computed with the quantile method, q=0.8. The largest \widehat{TDT} are obtained for the connection stations in the morning peak hours which is coherent with the commuting nature of line N.

due to the difference between alighting and boarding but also to the variability of the flows regarding passengers' behavior (e.g. more stressed passengers moving faster or boarding passengers in front of the doors). The attempted linear fit is of poor quality (having $R^2=0.44$) and it does not cross the origin. This is probably due to a change of regime for more than ten passengers' movements: $D\widehat{ABT}$ being more spread and increasing less fast for more than ten passengers' movements than for less.

652 8 CONCLUSIONS AND PERSPECTIVES

653 8.1 Main research findings

New passenger flow data with a higher resolution being accessible, we developed two new methods enabling us to compute the tight dwell time.

Contrary to existing methods, the cluster method and the quantile method enable an 656 extension of the situations where one can estimate a tight dwell time from only late stops 657 of late trains to any stop of any train. With these new methods, one can moreover estimate 658 the alighting and boarding time for any door. Compared to the quantile method, the 659 existing heuristics show weaknesses. Taking the observed dwell times of late trains is 660 an overestimation of the tight dwell time while Cornet's minimum dwell time does not 661 take into account enough variability. Indeed, the developed methods took into account the 662 instability of the passenger flow: for the same number of passengers alighting and boarding, 663 the necessary time to complete the alighting and boarding process may vary and change 664 the tight dwell time. Our methods capture this variability. 665



Figure 16: Average estimated tight dwell times (TDT) for all services on line N, branch Paris-Dreux toward Dreux computed with the quantile method, q=0.8. The largest \widehat{TDT} are obtained for the connection stations in the evening peak hours which is coherent with the commuting nature of line N.

666 8.2 Implications for practice

The extension of the estimation to any stop enables the computation of tight dwell times and margins for past circulations. Knowing those values will help to optimize the timetable for the future. Having an estimation for the alighting and boarding time at all the doors of the train gives an illustration of the distribution of the passengers' movements in the platform train interface. For example, some doors may have high *DABT* values while others may have low ones. Such an illustration would help to evaluate policies to make the alighting and boarding times more uniform among the doors.

674 8.3 Limitations of the current study

In this study, the assessment of the new methods we proposed was done by comparison 675 to existing methods. To ensure a fair comparison, we made the estimations in similar 676 subsets of the data (i.e. for late trains or non-early trains, see Table ??). This means 677 we have no assessment directly for non-late trains and the extrapolation of the results for 678 late trains should be done cautiously. Furthermore, the comparison we made was relative 679 to the existing methods. We highlighted that Cornet's minimum dwell time keeps little 680 margin but our methods may also have underestimated the tight dwell time (depending on 681 their parameter Δ_t or q) and we were not able to measure precisely to what extent. The 682 main issue there was that we did not compare the same object. One could always have 683 imagined a situation comparable where people alight and board faster which would have 684 given a smaller TDT. The methods we proposed compute the TDT as TDT = ABT + TT685 (TT being the technical time). This suggests that everything that is not included in the 686 alighting and boarding time is a margin and should be dropped to compute TDT. This 687 implies that the time for a particularly slow passenger to alight/board the train should 688 be considered as a margin (slow regarding the parameter Δ_t or q) and we know that it is 689 not always the case. Such an issue could be discussed in more detail using video of the 690



Figure 17: Mean door alighting and boarding times (DABT) 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. 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.



Figure 18: Door alighting and boarding times $(D\widehat{ABT})$ 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, $D\widehat{ABT}$ is not well modeled by a linear fit of passengers' movements.

⁶⁹¹ dwelling process but we only have access to passenger counting data.

692 8.4 Future research directions

As already mentioned, the methods we developed enabled us to analyze the tight dwell times for past circulations. Future research should focus on predicting TDT for future circulations. Knowing obtained \widehat{TDT} , some margin will have to be smartly added to build a robust timetable for an existing line. Some more advanced work will be able to make possible predictions of TDT for new missions, new stops, or even new lines.

698 **References**

Andersson, E., Peterson, A., & Krasemann Törnquist, J. (2011). Robustness in swedish railway
traffic timetables. In 4th international seminar on railway operations modelling and analysis.
Rome, Italy.

Andersson, E., Peterson, A., & Törnquist Krasemann, J. (2013). Quantifying railway timetable
 robustness in critical points. Journal of Rail Transport Planning Management, 3(3), 95-110. Re trieved from https://www.sciencedirect.com/science/article/pii/S2210970613000516
 (Robust Rescheduling and Capacity Use) doi: doi:10.1016/j.jrtpm.2013.12.002

Armstrong, J., & Preston, J. (2011). Alternative railway futures: growth and/or
 specialisation? Journal of Transport Geography, 19(6), 1570-1579. Retrieved
 from https://www.sciencedirect.com/science/article/pii/S0966692311001281
 doi: 10.1016/j.jtrangeo.2011.03.012

Brethomé, L. (2018). Modélisation et optimisation d'un plan de transport ferroviaire en zone dense
du point de vue des voyageurs (Doctoral thesis). Retrieved from https://www.theses.fr/
2018ECLI0014

Buchmueller, S., Weidmann, U., & Nash, A. (2008). Development of a dwell time calculation
model for timetable planning. WIT Transactions on the Built Environment, 103, 525–534. doi:
doi:10.2495/CR080511

Christoforou, Z., Chandakas, E., & Kaparias, I. (2020). Investigating the impact of dwell
time on the reliability of urban light rail operations. Urban Rail Transit, 6, 116–131. doi:
doi:10.1007/s40864-020-00128-1

- 719 Cornet, S., Buisson, C., Ramond, F., Bouvarel, P., & Rodriguez, J. (2019). Methods for quantita-
- tive assessment of passenger flow influence on train dwell time in dense traffic areas. Transporta-
- *tion Research Part C: Emerging Technologies*, 106, 345-359. doi: doi:10.1016/j.trc.2019.05.008
- Coulaud, R., Keribin, C., & Stoltz, G. (2023). Modeling dwell time in a data-rich railway environment: with operations and passenger flows data. *Transportation Research Part C: Emerging Technologies*, 146, 103980. doi: doi:10.1016/j.trc.2022.103980
- Daamen, W., Goverde, R. M., & Hansen, I. A. (2009). Non-discriminatory automatic registration
 of knock-on train delays. *Networks and Spatial Economics*, 9, 47–61.
- De Nailly, P. (2023). Analyse et prédiction des flux piétons dans un pôle de transport multi modal à partir de données multi-sources (Doctoral thesis). Retrieved from https://theses
 .hal.science/tel-04090167
- Dewilde, T., Sels, P., Cattrysse, D., & Vansteenwegen, P. (2011). Defining robustness of a railway
 timetable. In 4th international seminar on railway operations modelling and analysis. Rome,
 Italy.
- 733 DRIEA, O., Ile-de-France Mobilité. (2022). Enquête globale transport (egt) 2020.
- Givoni, M., Brand, C., & Watkiss, P. (2009). Are railways climate friendly? Built Environment,
 35(1), 70-86. Retrieved from https://www.ingentaconnect.com/content/alex/benv/2009/
 00000035/00000001/art00006 doi: doi:10.2148/benv.35.1.70
- Goverde, R. (2005). Punctuality of railway operations and timetable stability analysis (Doctoral the sis). Retrieved from 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
- 741 -stability-analysis.pdf
- Hansen, I. (2010). Railway network timetabling and dynamic traffic management. International Journal of Civil Engineering, 8. Retrieved from http://ijce.iust.ac.ir/article-1-422-en
 .html
- Harris, N. (2005). Train boarding and alighting rates at high passenger loads. Journal of Advanced
 Transportation, 40(3), 249–263. doi: doi:10.1002/atr.5670400302
- Harris, N. (2015). A european comparison of station stop delays. In International congress on (p. 1-6).
- IenW. (2022). Programma hoogfrequent spoorvervoer achtste oortgangsrapportage. Re trieved from https://www.rijksoverheid.nl/documenten/rapporten/2022/10/11/bijlage
 -1-phs-vgr-7-2022-1
- Kecman, P., & Goverde, R. M. (2013). Process mining of train describer event data and automatic
 conflict identification. Computers in Railways XIII: Computer System Design and Operation in
 the Railway and Other Transit Systems, 127, 227.
- Kecman, P., & Goverde, R. M. (2015). Predictive modelling of running and dwell times in railway
 traffic. *Public Transport*, 7, 295–319. doi: doi:10.1007/s12469-015-0106-7
- Kuipers, R. (2024). Understanding dwell times using automatic passenger count data: A quantile
 regression approach. Journal of Rail Transport Planning & Management, 29, 100431. doi:
 doi:10.1016/j.jrtpm.2024.100431
- Kuipers, R., & Palmqvist, C.-W. (2022). Passenger volumes and dwell times for commuter trains:
 A case study using automatic passenger count data in stockholm. *Applied Sciences*, 12. doi: doi:10.3390/app12125983
- Li, D., Daamen, W., & Goverde, R. M. (2016). Estimation of train dwell time at short stops
 based on track occupation event data: A study at a dutch railway station. Journal of Advanced
 Transportation, 50(5), 877–896.

- Luan, X., & Corman, F. (2022). Passenger-oriented traffic control for rail networks:
 An optimization model considering crowding effects on passenger choices and train operations. Transportation Research Part B: Methodological, 158, 239-272. Retrieved
 from https://www.sciencedirect.com/science/article/pii/S0191261522000273 doi:
 doi:10.1016/j.trb.2022.02.008
- ⁷⁷¹ McNally, M. G. (2007). The four step model. In *Handbook of transport modeling* (chap. 3).
- Medeossi, G., & Nash, A. (2020). Reducing delays on high-density railway lines:
 London-shenfield case study. *Transportation Research Record*, 2674(7), 193–205. doi: doi:10.1177/0361198120921159
- Mouwen, A., & Rietveld, P. (2013). Does competitive tendering improve customer satisfaction
 with public transport? a case study for the netherlands. *Transportation Research Part A: Policy*and Practice, 51, 29-45. doi: doi:10.1016/j.tra.2013.03.002
- Palmqvist, C.-W. (2019). Delays and timetabling for passenger trains (Doctoral thesis). Retrieved from http://portal.research.lu.se/ws/files/70626078/Carl_William Palmqvist_web.pdf
- Pedersen, T., Nygreen, T., & Lindfeldt, A. (2018). Analysis of temporal factors influencing
 minimum dwell time distributions. WIT Transactions on the Built Environment, 181, 447–58.
 doi: doi:10.2495/CR180401
- Puong, A. (2000). Dwell time model and analysis for the mbta red line. Massachusetts Institute
 of Technology Research Memo, 02139–4307.
- Seriani, S., & Fujiyama, T. (2019). Modelling the distribution of passengers waiting to board the
 train at metro stations. Journal of Rail Transport Planning Management, 11, 100141. doi:
 doi:10.1016/j.jrtpm.2019.100141
- van Loon, R., Rietveld, P., & Brons, M. (2011). Travel-time reliability impacts on rail way passenger demand: a revealed preference analysis. Journal of Transport Geography,
 19(4), 917-925. Retrieved from https://www.sciencedirect.com/science/article/pii/
 S0966692310001912 doi: doi:10.1016/j.jtrangeo.2010.11.009
- 793 Vromans, M. (2005). Reliability of railway systems (Doctoral thesis). Retrieved from https:// 794 repub.eur.nl/pub/6773/
- Weston, J. (1989). Train service model-technical guide. London underground operational research
 note, 89, 18.
- Wiggenraad, I. P. B. (2001). Alighting and boarding times of passengers at dutch railway stations
 (Doctoral thesis).
- Yamamura, A., Koresawa, M., Adachi, S., & Tomii, N. (2012). Identification of causes of delays in urban railways. WIT Transactions on The Built Environment, 127, 403–414. doi:
 doi:10.2495/CR120341
- Yang, J., Shiwakoti, N., & Tay, R. (2019). Train dwell time models development in the past
 forty years. In 41st australasian transport research forum (atrf). Canberra, Australia.
- Yap, M., & Cats, O. (2021). Predicting disruptions and their passenger delay impacts for public
 transport stops. *Transportation*, 48, 1703–1731. doi: doi:10.1007/s11116-020-10109-9

806 9 APPENDIX

807 9.1 Demonstration of the comparability of Cornet's \widehat{MDT} and estimated \widehat{TDT}

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

809
$$\lim_{|\{k,d\}|\to\infty} \left(\min_{k,d \ s.t. \ p=p_0} \widehat{M}_{k,s_0,d}\right) = 0$$
, then:

810
$$\lim_{|\{k,d\}|\to\infty} \left(\min_{k,d \ s.t. \ p=p_0} \widehat{TDT}_{k,s_0,d}\right) = \lim_{|\{k,d\}|\to\infty} \widehat{MDT}(p_0,s_0).$$

⁸¹¹ First, we trivially have (without any hypothesis):

812
$$\min_{k,d \ s.t. \ p=p_0} \widehat{TDT}_{k,s_0,d} \le \widehat{MDT}(p_0,s_0),$$

813 which holds if one takes the limit.

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

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

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

818
$$\lim_{|\{k,d\}|\to\infty} \left(\min_{k,d \ s.t. \ p=p_0} \widehat{M}_{k,s_0,d}\right) = 0$$

819 This implies that:

81

$$\widehat{ABT}_{k_0,s_0,d_0} + \widehat{M}_{k_0,s_0,d_0} \le \lim_{|\{k,d\}| \to \infty} \left(\min_{k,d \ s.t. \ p=p_0} \widehat{ABT}_{k,s_0,d} \right) + \epsilon.$$

821 In particular,

822
$$\left(\min_{k,d \ s.t. \ p=p_0} \widehat{ABT}_{k,s_0,d} + \widehat{M}_{k,s_0,d}\right) \leq \lim_{|\{k,d\}| \to \infty} \left(\min_{k,d \ s.t. \ p=p_0} \widehat{ABT}_{k,s_0,d}\right) + \epsilon.$$

⁸²³ This being true for any ϵ , one obtains:

$$\lim_{|\{k,d\}|\to\infty} \left(\min_{k,d\ s.t.\ p=p_0} \widehat{ABT}_{k,s_0,d} + \widehat{M}_{k,s_0,d}\right) \le \lim_{|\{k,d\}|\to\infty} \left(\min_{k,d\ s.t.\ p=p_0} \widehat{ABT}_{k,s_0,d}\right).$$

825 And, adding TT,

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

827 Therefore,

820

828

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

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