



## Prédire le temps de stationnement minimal à court terme : modèles et application

Rémi Coulaud<sup>a,b\*</sup>, Marc Deruelle<sup>a</sup>, Christine Keribin<sup>b</sup>, Pierre Messulam<sup>a</sup>, Gilles Stoltz<sup>b</sup>

*a Transilien, SNCF Mobilités, 10 rue Camille Moke, 93220, Saint-Denis, France*

*b Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, 91405, Orsay, France*

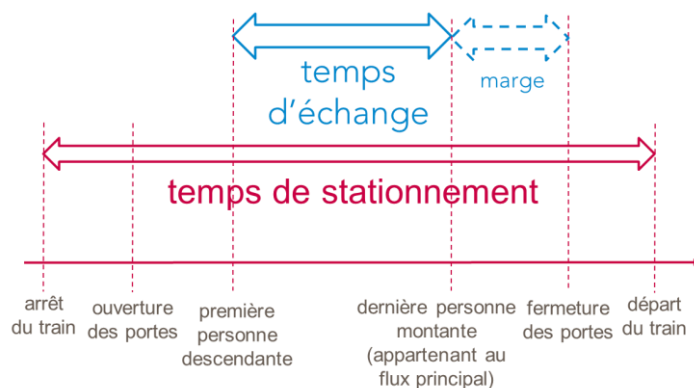
### A. Introduction

En Île de France, le temps de stationnement représente 25% à 30% du temps de trajet. En ce sens, la capacité d'une ligne ferroviaire ainsi que le temps de parcours des voyageurs sont directement impactés par le temps de stationnement.

Un opérateur ferroviaire cherche à réduire les excès de temps de stationnement tout en s'assurant qu'il soit suffisant pour éviter les effets « boule de neige » grâce aux marges opérationnelles (Van Breusegem et al., 1991, Hansen et al. 2009). Aujourd'hui, la SNCF utilise le temps de stationnement comme un amortisseur d'aléas. Un travail récent de Palmqvist et al. (2017) illustre ce compromis entre précision et marge que doit trouver un opérateur ferroviaire.

On rappelle que le temps de stationnement est l'intervalle de temps passé à quai entre l'heure d'arrivée et de départ du train. Ce dernier se décompose en plusieurs sous parties (Buchmüller et al., 2008) dont le temps technique, le temps nécessaire à l'ouverture et à la fermeture des portes et le temps d'échange, temps nécessaire à l'échange des flux de montées et de descentes, voir figure 1. En laboratoire ou par simulation multi-agents de nombreux auteurs ont montré de façon

analytique que le temps d'échange pouvait être une fonction linéaire (Daamen et al., 2008) ou non linéaire (Yamamura et al., 2013) du nombre de montées et de descentes par porte.



La SNCF ne mesure pas le temps d'échange mais elle mesure très précisément le temps de stationnement. Il nous est impossible de déduire directement

Figure 1 : décomposition du temps de stationnement

\* Auteur correspondant  
Adresse mail: remi.coulaud@math.u-psud.fr



du temps de stationnement le temps d'échange étant donné que le plan de transport contraint l'heure d'arrivée et de départ du train à l'inverse des « short stop » à la néerlandaise (Li et al., 2016), où le conducteur ferme les portes dès que l'échange est terminé.

Cependant à la SNCF, si le train est en retard, le conducteur doit rattraper le retard en démarrant dès l'échange terminé. Ainsi, nous nous focalisons sur le temps de stationnement minimal défini par Pedersen et al. (2018) : c'est-à-dire le temps de stationnement pour les trains en retard à l'arrivée qui ne sont alors plus contraints par le plan de transport théorique. Cette idée est utilisée par Hansen et al. (2009) afin de mesurer la durée de trajet minimale entre deux gares. Comme le remarquent justement Cornet et al. (2019), rien ne garantit en pratique que la voie soit libre ni que le conducteur respecte scrupuleusement la règle en cas de retard.

Notre objectif est de prédire par train le temps de stationnement minimal en fonction des conditions d'exploitation afin de connaître les marges opérationnelles existantes. Le modèle souhaité donne à chaque arrêt le temps de stationnement optimal à un horizon d'un arrêt étant donné le train (nombres de places disponibles), la gare (distribution à quai des voyageurs liée à la géométrie des quais actuel et à venir), les autres trains (écart temps entre les trains). A terme, nous souhaitons obtenir pour tous les arrêts d'un train.

Nos contributions sont les suivantes :

1. Comparaison de modèles de prévision du temps de stationnement à court terme dans le cas de trains en avance, à l'heure ou en retard par rapport à l'heure d'arrivée
2. Utilisation de l'effet de la distribution à quai pour prédire le temps de stationnement en lien avec les travaux initiés par (Wiggenraad, 2001)

## **B. Méthodologie**

Nous partons des données collectées par Transilien pour prédire le temps de stationnement minimal pour chaque train. Nous proposons un modèle prédictif pour le temps de stationnement dans trois situations : trains en avance, à l'heure et en retard. C'est le dernier cas qui nous donne accès au temps de stationnement minimal. Nous proposons d'évaluer les performances de prédiction de modèles connus de la littérature, de modèles naïfs ainsi que de nos propres modèles.

### **1. Données**

Nous avons accès à plusieurs sources de données d'août 2017 à avril 2019 pour la ligne Transilien H.



## MONTANTS / GARE

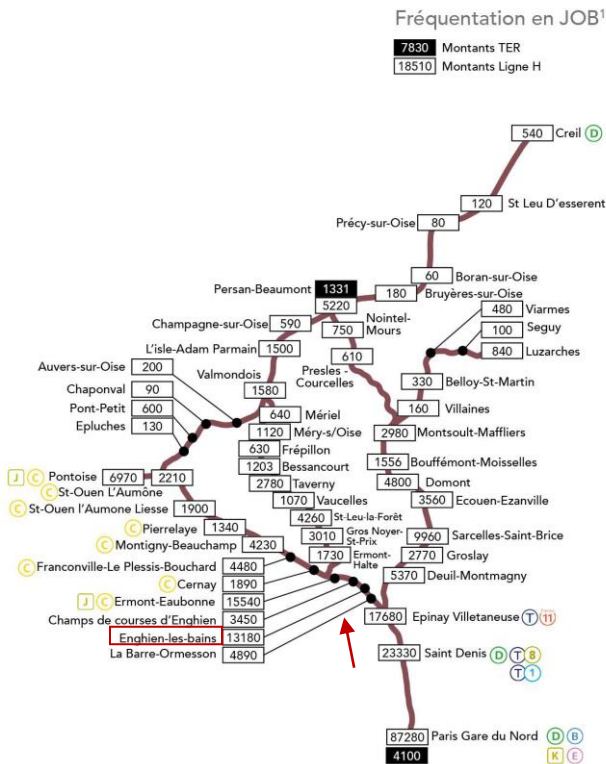


Figure 2 : ligne H

temporelle qui sépare le train arrivé à quai des derniers départs sur la même branche. Ces données sont pour le moment accessibles à J+2.

Pour chaque gare  $s$  et chaque train  $k$ , nous avons à disposition les variables aléatoires décrites dans le tableau 1.

	Variables	Disponibilité	Notation
Variable à prédire	Temps de stationnement (s)	J+2	$Y_s^k$
Variables de fréquentation	Montées, descentes	J+2	$M_s^k, D_s^k$
	Taux d'occupation	J+2	$TO_s^k$
Variables d'exploitation ferroviaire	Temps de stationnement théorique (s)	Instantané	$YTH_s^k$
	Ecart à l'heure d'arrivée théorique (s)	J+2	$\Delta ATH_s^k$
	Ecart au train précédent (s)	J+2	$\Delta TT_s^k$
	Temps de stationnement à la gare précédente (s)	J+2	$Y_{s-1}^k$
	Temps de stationnement du train précédent à la même gare précédente (s)	J+2	$Y_s^{k-1}$

Tableau 1 : variables à disposition

Nous nous concentrons sur la gare d'Enghien-les-Bains pour les trains impairs (Paris-Banlieue), voir figure 2. Enfin, nous prévoyons le temps de stationnement uniquement pour les trains circulant en heures de pointe en soirée, estimées à partir d'un mélange gaussien, pour les jours du lundi au vendredi hors jours fériés et vacances.

La première mesure utilisée est celle du temps de stationnement à partir de l'odométrie des rames z50000. La deuxième mesure provient de l'APC (Automatic Passenger Counting) pour chaque rame de 8 voitures (une voiture = une porte) les entrées et sorties à chaque porte sont comptées. Nous avons aussi accès à la charge à bord à partir d'une somme cumulée.

Nous croisons ces données avec le plan de transport théorique pour obtenir l'heure d'arrivée et de départ théoriques. Nous calculons la distance



Pour la prévision, nous supposons connu l'historique complet d'un train et les flux par porte. Ainsi, nous avons à disposition la réalisation des variables aléatoires décrites dans le tableau 1 pour environ deux ans de circulation :

	Entrainement	Test
Nombre d'observations : gare d'Enghien-les-Bains	2558	729

## 2. Critères de comparaison

L'objectif est de prévoir le temps de stationnement pour trois situations à partir d'un ensemble de variables explicatives, nous mesurons l'erreur commise sur le jeu de données test constitué par le début de l'année 2019.

Trois types d'erreurs sont calculées : l'erreur moyenne quadratique (RMSE), l'erreur moyenne absolue (MAE) et l'erreur absolue relative (MAPE).

## 3. Modèles existants

À notre connaissance aucun article ne compare les performances de prédiction des modèles statistiques existant sur un même jeu de données : certains auteurs utilisent plusieurs jeux de données pour évaluer une même méthode (Li et al., 2016, Palmqvist et al., 2019), d'autres font un état de l'art des méthodes existantes (Li et al., 2016, Cornet et al., 2019). Nous proposons un bref récapitulatif des modèles existants, voir tableau 2, que nous utilisons sur notre jeu de données. Nous reconstruisons les résultats de trois modèles connus de la littérature :

Article	Modèle	Adaptation
(Kecman et al., 2015)	Régression robuste (LTS) avec le package R robustbase (Rousseuw et al., 2014) Forêt aléatoire avec le package R randomForest (Liaw et al., 2002)	Nous considérons un modèle simplifié qui ne prend en compte que l'écart à l'heure d'arrivée théorique et le temps de stationnement théorique
(Li et al., 2016)	$Y_s^k = \beta_0 + \beta_1 L_s^k + \beta_2 L_s^{k-1} + \beta_3 Y_s^{k-1} + \beta_4 \sqrt{Y_{s-1}^k Y_{s-2}^k}$	Nous n'utilisons pas la longueur du train $L_s^k$ mais la composition du train
(Palmqvist et al., 2019)	$Y_s^k = \beta_0 + \beta_1 M_s^k + \beta_2 D_s^k + \beta_3 TO_s^k$	Nous reconstruisons le modèle complet avec les variables quadratiques et d'interactions

Tableau 2 : modèles existants

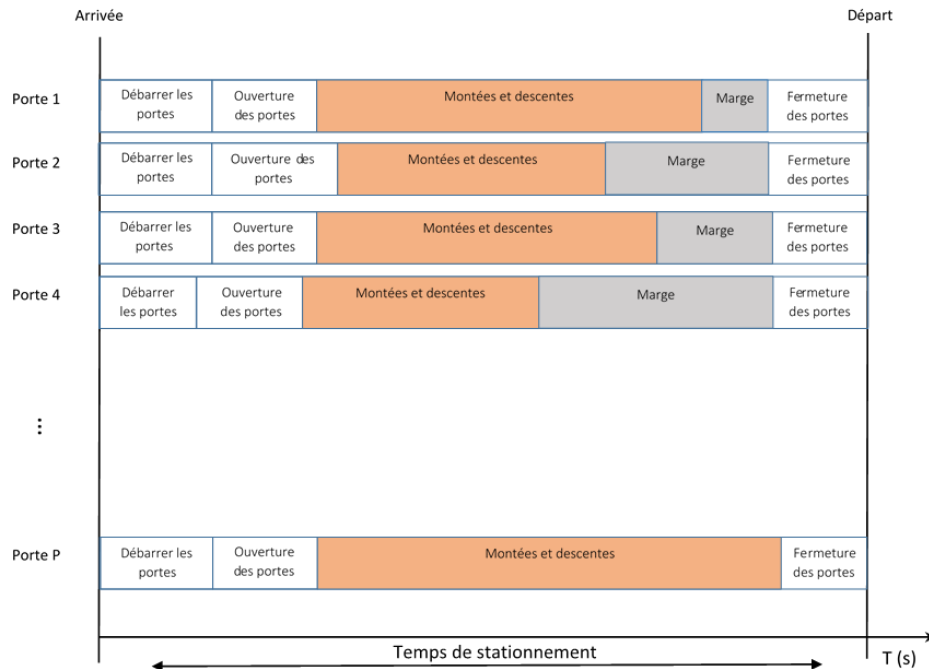


Figure 3 : parallélisations de l'échange voyageur à quai

#### 4. Modèles naïf et théorique

Le modèle naïf consiste à prédire le temps de stationnement du train à la gare  $s$  à partir de la moyenne des temps de stationnement observés pour le même train  $k$  dans le passé.

Le modèle théorique consiste à utiliser le temps de stationnement programmé dans le plan de transport pour prédire le temps de stationnement. C'est un moyen d'évaluer la pertinence du temps de stationnement théorique calculé par la SNCF.

#### 5. Modèle proposé : une rupture avec l'hypothèse d'uniformité des échanges voyageurs le long d'un train

Nous proposons un modèle qui prend en compte le fait que le nombre de montées et descentes ne se fait pas de façon uniforme entre les différentes portes, voir figure 3 inspirée de Buchmüller et al. (2008). Les portes proches des entrées ou sorties de quais seront davantage utilisées que des portes excentrées. Cette intuition, bien que déjà présente chez Lin et al. (1992) a été écartée pour des raisons techniques. Nous définissons l'affluence critique comme :

$$AC_s^k = \max_{p \in \{1, \dots, P\}} A_s^{k,p}$$

Avec l'affluence pour chaque porte  $p$  de 1 à  $P$  définie par :

$$A_s^{k,p} = M_s^{k,p} + D_s^{k,p}$$



L'affluence critique pour l'exemple de la figure 3 est celle de la porte P. Nous utilisons une régression linéaire classique ainsi qu'une forêt aléatoire comme Kecman et al. (2015).

Nous	$Y_s^k = \beta_0 + \beta_1 M_s^k + \beta_2 D_s^k + \beta_3 YTH_s^k + \beta_4 AC_s^k + \beta_5 TO_s^k + \beta_6 \Delta ATH_s^k + \beta_7 \Delta TT_s^k + \beta_8 Y_s^{k-1} + \beta_9 Y_{s-1}^k$	...	...
------	--	-----	-----

### C. Résultats attendus

Les résultats attendus sont tout d'abord une meilleure prévision du temps de stationnement des modèles utilisant les variables ferroviaires pour les trains en avance. Tandis que les variables de fréquentation permettent de mieux prédire le temps de stationnement lorsque ce dernier est tendu notamment à cause de retard.

De plus, les temps de stationnement pour les trains en avance étant plus long nous nous attendons à ce que les erreurs absolues et quadratiques soient plus grandes dans ce cas.

L'affluence critique permet de mieux comprendre/prédire le temps de stationnement.

Enfin, nous montrons l'importance des marges dans la construction des temps de stationnement théoriques à la SNCF.



#### D. Bibliographie

1. Van Breusegem, V., Campion, G. & Bastin, G. (1991). Traffic modeling and state feedback control for metro lines. *IEEE Transactions on Automatic Control* **36**, 770–784.
2. Hansen, I. A. (2010). Railway Network Timetabling and Dynamic Traffic Management. *International Journal of Civil Engineering* **8**, 19–32.
3. Kecman, P. & Goverde, R. M. P. (2015). Predictive modelling of running and dwell times in railway traffic. *Public Transport* **7**, 295–319.
4. Cornet, S., Buisson, C., Ramond, F., Bouvarel, P. & Rodriguez, J. (2019). Methods for quantitative assessment of passenger flow influence on train dwell time in dense traffic areas.
5. Li, D., Daamen, W. & Goverde, R. M. P. 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**, 877–896 (2016).
6. Lin, T. & Wilson, N. H. M. (1992). Dwell time relationships for light rail systems. *Transportation Research Record*.
7. Buchmüller, 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.
8. Daamen, W., Lee, Y. & Wiggeraad, P. (2008). Boarding and Alighting Experiments: Overview of Setup and Performance and Some Preliminary Results. *Transportation Research Record*.
9. Pedersen, T., Nygreen, T. & Lindfeldt, A. (2018). Analysis of temporal factors influencing minimum dwell time distributions. *WIT Transactions on The Built Environment*. **181**, 447–458
10. Palmqvist, C.-W., Olsson, N. & Hiselius, L. (2017). An Empirical Study of Timetable Strategies and Their Effects on Punctuality.
11. Wiggeraad, P. (2001). Alighting and boarding times of passengers at Dutch railway stations-analysis of data collected at 7 stations in October 2000. *TRAIL Research School: Delft University of Technology, Delft*.



### **Mots clés**

Temps de stationnement minimal, modélisation statistique, prévision, flux voyageur-ses, distribution à quai

### **Sessions visées**

1. Session n° 2 : Gestion du trafic, systèmes de transport intelligent, management des infrastructures et des réseaux
2. Session n° 27 : Flux de voyageur-ses : quelle gestion dans les gares ?
3. Session n° 33 : Valorisation des données de transport collectif