

Table des matières

	Page
Introduction	1
1 Éléments de trafic – problématique	5
1.1 Le chemin de fer, matériel roulant et infrastructure	6
1.1.1 Le matériel roulant	6
1.1.2 L'infrastructure	11
1.2 Le chemin de fer : gestion des circulations	20
1.2.1 Planification : gestion à l'horaire ou à la fréquence	20
1.2.2 Les sillons	21
1.2.3 Débit et capacité	22
1.2.4 Construction du plan de transport et robustesse	24
1.2.5 Effondrement du plan de transport	28
1.3 Le trafic routier et la congestion	30
1.3.1 Variables individuelles-variables de flux	30
1.3.2 Relation entre vitesse, concentration et débit du flux.	33
1.3.3 Diagramme fondamental	33
1.3.4 Le modèle LWR	34
1.3.5 Congestion-définitions	37
1.3.6 Comparaison entre le trafic ferroviaire et le trafic routier	37
1.4 Problématique	41
2 Modélisation microscopique stochastique sans gare	43
2.1 État de l'art de la modélisation stochastique du déplacement des trains	44
2.1.1 Historique de la modélisation stochastique du déplacement des trains	44
2.2 Construction du modèle	49
2.2.1 L'infrastructure	49
2.2.2 Le matériel roulant	49
2.2.3 Les forces décrivant le comportement d'un train	49
2.2.4 Implémentation numérique du modèle	54
2.3 Apparition de la congestion	55
2.3.1 Organisation des simulations	55
2.3.2 Compétition entre la force aléatoire et la force horaire	56
2.3.3 Résultats	58
2.4 Analyse du modèle et des résultats par des processus de Markov	62
2.4.1 Processus de Markov	62

2.4.2	Les distributions de type phase	63
2.4.3	Lien avec le modèle	68
2.5	Quantification de la congestion	73
2.5.1	Organisation des simulations	73
2.5.2	Dynamique de la vitesse instantanée moyenné sur les trains	73
2.5.3	Diagramme Fondamental de Ligne Ferroviaire	75
3	Prise en compte des gares	81
3.1	Évolution du modèle pour prendre en compte les gares	82
3.1.1	L'infrastructure et le plan de transport	82
3.1.2	Implémentation numérique du modèle	88
3.2	Apparition de la congestion	90
3.2.1	Organisation des simulations	90
3.2.2	Résultats	91
3.3	Quantification de la congestion	98
3.3.1	Organisation des simulations	98
3.3.2	Dynamique de la vitesse moyenne instantanée	98
3.3.3	Diagramme Fondamental de Ligne Ferroviaire	101
3.4	Conclusions sur le modèle développé	104
3.4.1	Aspects statiques	104
3.4.2	Aspects dynamiques	105
4	Confrontation avec les données	109
4.1	Données	110
4.2	Impossibilité d'un diagramme fondamental local	112
4.3	Vers un diagramme fondamental de ligne ferroviaire	113
4.3.1	Principes du diagramme fondamental de réseau en routier	113
4.3.2	Méthode de construction du DFLF	114
4.3.3	Résultats et analyses	115
4.4	Confrontation des données avec le modèle	119
4.4.1	Diagramme fondamental de ligne ferroviaire moyenné : cas réel	119
4.4.2	Mise en perspective avec la modélisation	120
4.5	Discussion et conclusion	122
Conclusion		125
Bibliographie		132
A Détermination des paramètres de la force horaire		1
B Détermination des paramètres de la force aléatoire		4
C compétition entre la force aléatoire et la force horaire		6
D Calcul d'optimisation de la taille des cantons		8

Liste des figures

	Page
1.1 Bilan des forces physiques	7
1.2 Graphe des composantes de la force appliquée au train	10
1.3 Exemple de signalisation à trois aspects	17
1.4 Comparaison de réseaux urbains routier et ferroviaire	18
1.5 plan du réseau régional parisien	19
1.6 graphique de circulation	22
1.7 distance minimale prévue entre deux trains	23
1.8 détail d'un graphique de circulation	26
1.9 trajectoires de véhicules et zone d'Édie	31
1.10 Exemples de diagramme fondamental expérimental sur autoroute	35
1.11 Diagramme fondamental triangulaire	36
2.1 Graphe des composantes de la force appliquée au train	50
2.2 Bornes de l'accélération d'un train	51
2.3 Espace des paramètres exploré	58
2.4 Probabilité d'apparition de la congestion au cours du temps	60
2.5 Exemples de distributions des écarts horaires primaires	61
2.6 Exemple d'interpolation exponentielle	64
2.7 Exemple de représentations des DPT	65
2.8 Représentation du graphe et de la matrice de notre DTP	66
2.9 Représentation de $s_i = f(i)$ pour différents cas de figure	67
2.10 Erreur quadratique moyenne en fonction de k	69
2.11 Exemple d'interpolation DTP	70
2.12 Évolution de s_9 en fonction des paramètres du modèle	71
2.13 Illustration des distances limites entre deux trains	71
2.14 Limite d'apparition de la congestion	72
2.15 $\bar{v}^{T,500}(t)$	76
2.16 Diminution de la vitesse moyenne en congestion	77
2.17 DFLF	78
3.1 Planification de la marche d'un train entre deux gares	84
3.2 Position des cantons et marche du train	89
3.3 Espace des paramètres explorés	92
3.4 Probabilité d'apparition de la congestion au cours du temps	94
3.5 Exemples d'interpolations DTP	95

3.6	$\bar{v}^{T,500}(t)$	99
3.7	Diminution de la vitesse moyenne en congestion	100
3.8	DFLF avec gares, variations de la force aléatoire	102
3.9	DFLF avec gares, variations du temps d'arrêt	103
3.10	Effets des aléas sur le DFLF	105
4.1	Portion du réseau analysé	111
4.2	Calendrier des jours analysés	112
4.3	Diagramme fondamental expérimental entre deux balises	113
4.4	Exemples de DFLF	116
4.5	Exemples de DFLF avec interpolation	118
4.6	Analyse d'une chute de débit	118
4.7	DFLF moyennés sur l'ensemble des jours	121
4.8	Effets des aléas sur le DFLF	122
B.1	distribution de F_{rand}	4
D.1	Position des cantons et marche du train	8

Liste des tableaux

1.1	comparatif ferroviaire-routier	40
2.1	Chronologie du stochastique en ferroviaire	45
2.2	Notations	48
3.1	Descriptpion du motif de base de la simulation	89
3.2	Notations	90
4.1	Moyennes des résidus des régressions	117

Liste des publications

Publications dans une revue à comité de lecture :

CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEBOUL et D. DE ALMEIDA. 2015, «Analyzing railroad congestion in a dense urban network through the use of a road traffic network fundamental diagram concept», *Public Transport*, p. 1–13.

CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEBOUL et D. DE ALMEIDA. (en révision), «Global formulation of the transition probability of a theoretical railroad line to a congested state», *Transportation Research part B*.

Publications dans des actes de congrès à comité de sélection ou de lecture :

CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEBOUL et D. DE ALMEIDA. 2013, «Analysing railroad congestion in a dense urban network through the use of road traffic network fundamental diagram concept», dans *Models and Technologies for Intelligent Transportation Systems (MT-ITS)*.

CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEBOUL et D. DE ALMEIDA. 2014a, «Analysing railroad congestion in a dense urban network through the use of road traffic network fundamental diagram concept», dans *Transportation Research Board 93rd Annual Meeting*.

CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEBOUL et D. D. DE ALMEIDA. 2014b, «Analysing railroad congestion in a dense urban network through the use of road traffic network fundamental diagram concept», dans *Transport Research Arena (TRA) 5th Conference : Transport Solutions from Research to Deployment*.

CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEBOUL et D. D. DE ALMEIDA. 2015c, «Modelling the onset of congestion due to stochastic behavior of the components of the railroads system», dans *EURO Working Group on Transportation (EWGT)*.

Introduction

Contexte de la recherche

Depuis une vingtaine d'années, les problématiques de transport public en région parisienne sont devenues une préoccupation majeure, tant pour les pouvoirs publics que pour les usagers. Ces préoccupations s'inscrivent dans un contexte tendu : en regroupant à elle seule presque 20% de la population française, la région Ile-de-France doit satisfaire une très forte demande de transport. Cette dernière est renforcée par une organisation spatiale caractérisée par une forte dissociation entre les lieux de domicile et de travail. Le développement d'importants bassins d'emploi dans des zones périphériques contribue aujourd'hui à une accélération de cette tendance. Pour les salariés franciliens, qui consacrent en moyenne environ deux heures quotidiennement à leurs déplacements domicile-travail, ainsi la qualité de l'offre de transport est un enjeu majeur.

Le développement des transports publics représente aussi un double enjeu écologique et de santé publique dans une région où la voiture d'une part comptabilise plus de la moitié des déplacements et d'autre part contribue de manière significative aux émissions de CO₂ et à la dégradation de la qualité de l'air. Les transports publics, et tout particulièrement le ferroviaire, ont une empreinte écologique bien plus faible que la voiture individuelle, ce qui en fait des outils privilégiés pour relever les défis écologiques et sanitaires.

Deux compagnies se partage le secteur ferroviaire en Ile-de-France : La RATP¹ qui exploite le métro mais aussi deux lignes du RER², et Transilien (l'entité de la SNCF³ en charge des transports en région parisienne) qui exploite 14 lignes de RER ou de trains de banlieue. Ces acteurs jouent un rôle central dans l'organisation des transports en commun de cette région : le réseau ferré structure les espaces en définissant les grands axes de mobilité autour desquels se développent des réseaux de transports, souvent de bus, qui permettent une desserte plus locale. Mais, à l'opposé des attentes qui pèsent sur ce secteur, les RER et les trains de banlieue rencontrent un certain nombre de dysfonctionnements. Saturation des trains aux heures de pointes, retards à répétition et autres incidents font partie du quotidien de bon nombre d'usagers de ces réseaux de transports.

Les pouvoirs publics, la RATP et Transilien ont pris la mesure de ces problèmes et œuvrent, depuis une dizaine d'années, de façon concertée, à l'amélioration de l'état du

1. Régie Autonome des Transports Parisiens

2. Réseau Express Régional

3. Société Nationale des Chemins de fer Français

transport ferroviaire en région parisienne. Dans ce contexte, Transilien a entamé une démarche globale de remise en question de ses principes d'exploitation et s'est retrouvé, entre autres, confronté au besoin de mieux comprendre les phénomènes à l'origine des retards qui affectent ses trains. Il semble en effet que l'intensification rapide du trafic ferroviaire en Ile-de-France au cours des 40 dernières années ait montré les limites des méthodes et des équipements ferroviaires traditionnels. L'accélération du renouvellement des équipements ferroviaires fait partie des solutions retenues pour préparer l'avenir, mais Transilien ne peut se reposer sur cette seule démarche. Une réflexion est ouverte sur la possibilité de développer de nouvelles méthodes d'exploitation dites de «zone dense» en opposition aux méthodes historiques, par ailleurs parfaitement opérationnelles lorsque le trafic est peu dense.

L'idée de ce travail de thèse est de répondre au besoin de Transilien d'identifier les mécanismes qui en zone dense sont responsables de l'apparition des retards. La question originale et originelle de cette thèse a été de se demander si la théorie du trafic routier et les méthodes d'observation et d'analyse associées et qui ont fait leur preuve en trafic routier, pouvaient inspirer une approche innovante du trafic ferroviaire en zone dense qui réponde à un tel besoin.

Dans cette thèse, nous allons étudier la faisabilité de la transposition de deux outils du trafic routier vers le trafic ferroviaire : le diagramme fondamental local et le diagramme fondamental de réseau. On va mettre en évidence l'impossibilité du premier en ferroviaire, pour construire le second à partir de données expérimentales. Ce dernier va nous permettre d'observer et de quantifier l'état dégradé du trafic. Les résultats obtenus seront également comparé avec ceux issus d'un modèle simplifié permettant de reproduire l'état du trafic dégradé sous l'influence de perturbations aléatoires.

Plan de la thèse

Les travaux de thèse s'articulent autour de quatre chapitres :

- Le chapitre 1 est l'occasion de présenter d'une part le fonctionnement du système ferroviaire, d'autre part les bases de la théorie du trafic routier. Ces deux présentations sont l'occasion de faire ressortir les différences entre ces deux modes de transport afin d'orienter la problématique scientifique de cette thèse.
- Le chapitre 2 présente la construction d'un modèle simplifié du trafic ferroviaire en l'absence de gare. Ce modèle fait l'hypothèse que de petits événements viennent perturber le plan de transport et qu'on peut les modéliser de manière stochastique. On verra que sous certaines conditions ces événements peuvent se propager et impacter plusieurs trains. On étudiera l'aspect dynamique mais aussi statique de l'impact de ces perturbations, notamment à travers un outil d'analyse directement inspiré du trafic routier.
- Le chapitre 3 sera l'occasion de revenir sur certaines simplifications de ce modèle en introduisant des gares, une certaine hétérogénéité de l'infrastructure et en diver-

sifiant la façon de modéliser les perturbations stochastiques. Avec cette version du modèle, on caractérisera aussi les aspects dynamiques et statiques de l'impact de ces perturbations.

- Le chapitre 4 sera l'occasion de confronter nos modèles avec des données issues des lignes B et C du réseau Transilien. On présentera une méthode de construction d'un diagramme fondamental adapté en général au trafic ferroviaire et en particulier aux données disponibles. À l'aide de cet outil, on présentera une analyse des données disponibles. Les résultats seront ensuite comparés avec ceux des chapitres 2 et 3.

Chapitre 1

Éléments de trafic ferroviaire et routier – problématique de la thèse

Sommaire

1.1 Le chemin de fer, matériel roulant et infrastructure	6
1.1.1 Le matériel roulant	6
1.1.2 L'infrastructure	11
1.2 Le chemin de fer : gestion des circulations	20
1.2.1 Planification : gestion à l'horaire ou à la fréquence	20
1.2.2 Les sillons	21
1.2.3 Débit et capacité	22
1.2.4 Construction du plan de transport et robustesse	24
1.2.5 Effondrement du plan de transport	28
1.3 Le trafic routier et la congestion	30
1.3.1 Variables individuelles-variables de flux	30
1.3.2 Relation entre vitesse, concentration et débit du flux.	33
1.3.3 Diagramme fondamental	33
1.3.4 Le modèle LWR	34
1.3.5 Congestion-définitions	37
1.3.6 Comparaison entre le trafic ferroviaire et le trafic routier	37
1.4 Problématique	41

1.1 Le chemin de fer, matériel roulant et infrastructure

Définition - Chemin de fer

Le chemin de fer est un système de transport guidé destiné à l'acheminement de personnes ou de marchandises qui comporte :

- Une infrastructure qui comprend à minima :
 - Un ensemble de rails en acier qui forme la voie ferrée et assurent la fonction de guidage des véhicules.
 - Un ensemble de points d'arrêt, appelés gares ou stations, qui assurent la fonction d'interface d'échange entre deux véhicules ou entre un véhicule et l'extérieur du système.
- Un ensemble de véhicules de transport qui se déplace au moyen de roues en contact avec le rail.

1.1.1 Le matériel roulant

Les véhicules ferroviaires s'organisent en convois que l'on appelle des trains. Un train peut être composé d'une rame unique et parlera d'unité simple ou composé d'un attelage de plusieurs rames et on parlera d'unité multiple. Une rame peut être :

- Un attelage de voitures non motrices pour le transport de passagers avec une locomotive placée à une ou deux extrémités de la rame et qui assure la motorisation de la rame¹.
- Un attelage de wagons pour le transport des marchandises avec une locomotive placée à une ou deux extrémités de la rame et qui assure la motorisation de la rame.
- Un attelage d'éléments automoteurs qui sert pour le transport de passagers. Chaque élément étant moteur, il n'y a pas de locomotive mais une cabine de conduite aux deux extrémités de la rame.

Dans cette thèse, on se place dans le cadre du transport de passagers uniquement. On décrira souvent les trains sans rentrer dans le détail de leur composition mais en les résumant par un ensemble de caractéristiques physiques nécessaires et suffisantes pour décrire la dynamique de ce train avec une précision adaptée aux problématiques de cette thèse.

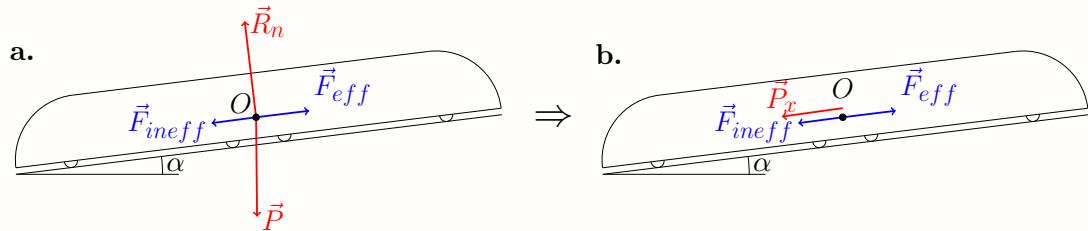
Modèle de la dynamique d'un train

Considérons un train comme un solide indéformable se déplaçant sur une voie qui fait un angle α avec l'horizontale². On fait le bilan des forces qui s'appliquent à celui-ci dans un cas général (cf. figure 1.1). On distingue quatre forces :

1. Dans certains trains longs de frets il peut y avoir une locomotive supplémentaire au milieu.
 2. Dans le langage ferroviaire on appelle cela une rampe en montée ou une pente en descente.

- Le poids \vec{P} qui est vertical et dirigé vers le bas et qui s'applique dans l'hypothèse d'un solide indéformable au centre de gravité du train.
- La réaction normale \vec{R}_n du rail à l'écrasement du train. Cette force s'applique aux points de contacts entre les roues et le rail perpendiculairement à celui-ci.
- Les frottements inefficaces \vec{F}_{ineff} dus à la résistance de l'air, aux zones de glissements entre les roues et le rail et à la résistance à l'arrachement. Ces frottements sont indésirables.
- Les frottements efficaces \vec{F}_{eff} qui transmettent les efforts de traction et de freinage au train par les surfaces où les roues (motrices pour les efforts de traction) adhèrent aux rails. Ces frottements sont utiles.

Figure 1.1 - Bilan des forces physiques



Bilan des forces physiques appliquées à un train. Bien que ce ne soit pas le cas en réalité, on positionne le point d'application de toutes les forces en un même point O. On distingue sur la figure **a** : le poids \vec{P} , la réaction normale \vec{R}_n , les frottements inefficaces \vec{F}_{ineff} et les frottements efficaces \vec{F}_{eff} . Ces derniers correspondent ici par choix à une force motrice et sont donc dirigés vers la droite. Sur la figure **b** on présente \vec{P}_x la résultante du poids et de la réaction normale afin que toutes les forces soient parallèles au rail.

La réaction normale n'existe que si le train est en contact avec le rail et elle représente l'impossibilité pour le train de pénétrer dans le rail, les seuls mouvements possibles se font donc sur le rail. Il en résulte que la somme du poids et de la réaction normale est parallèle au rail. L'axe formé par le rail est doté d'une abscisse x . La somme $\vec{P} + \vec{R}_n$ est aussi la projection de \vec{P} sur l'axe x , on l'appelle \vec{P}_x , et on la nomme force de déclivité. Ainsi il ne reste que trois forces \vec{P}_x , \vec{F}_{eff} et \vec{F}_{ineff} qui sont toutes parallèles au rail, on retombe sur un problème de mécanique à une dimension.

La force de déclivité \vec{P}_x s'exprime simplement comme la projection du poids sur l'axe x , son expression est présentée dans le formulaire 1.1.

La valeur de la force de frottement efficace \vec{F}_{eff} est limitée par la condition d'adhérence qui s'écrit selon la loi empirique de Coulomb :

$$\begin{aligned}\|\vec{F}_{eff}\| &\leq fac. \cdot \|\vec{R}_n\| = fac. \cdot m \cdot g \cdot \cos(\alpha) \\ &= fac. \cdot m \cdot g\end{aligned}$$

Avec $fac.$ le coefficient d'adhérence entre deux surfaces en acier, m le poids du train, g l'accélération de la gravité et $\alpha \ll 1$. Dans le cas du freinage, le seul facteur limitant est la

condition d'adhérence. Dans le cas de la traction, la condition d'adhérence et la puissance motrice du train sont toutes les deux des facteurs limitants, et on fait par ailleurs l'hypothèse simplificatrice que la puissance motrice du train est indépendante de la vitesse. Les formules correspondantes sont présentées dans le formulaire 1.1.

La force de frottement inefficace se décompose en plusieurs composantes :

- La force d'arrachement F_{ar} c'est la force minimale à fournir pour mettre le train en mouvement, une fois le train mobile cette force disparaît.
- Les frottements induits par la courbure de la voie F_{courbe} qui dépendent de l'écartement des voies et du rayon de courbure.
- Les frottements permanents F_{perm} dans lesquels on inclut également la résistance au roulement bien que ce ne soit pas à proprement parler un phénomène de frottement.
- Les frottements dus aux défauts $F_{défaut}$. Cette force dépend des défauts présents sur la voie et les roues et de la vitesse du train.
- Les frottements aérodynamiques F_{air} qui dépendent de la vitesse du train (on fait l'hypothèse de l'absence de vent).

Toutes ces composantes font intervenir des mécanismes complexes et sont en pratique déterminées de manière empirique ou semi-empirique. Les formules présentées dans le formulaire 1.1 sont issues de [ALLENBACH et collab. \[2008\]](#); [LAMALLE \[1947\]](#), les changements d'unité demandent quelques calculs d'adaptations. Il faut retenir davantage l'ordre de grandeur de ces coefficients qu'une valeur exacte. D'une part parce que ces coefficients varient beaucoup en fonction de la diversité des situations rencontrées, d'autre part parce que les formules présentées sont elles aussi imprécises. Il est important de noter que ces ordres de grandeurs correspondent à des roues en acier et qu'un train sur pneumatiques n'aurait pas le même comportement. Les valeurs numériques des paramètres fournis correspondent aux ordres de grandeurs observés pour un train en unité simple sur le réseau régional d'Île-de-France.

Conséquences de la faible adhérence

Les phénomènes de frottement et de résistance au roulement dus aux déformations qui ont lieu entre la roue et le rail sont responsables de l'adhérence du train mais occasionnent également des pertes d'énergie. Il n'existe pas de matériau qui assure à la fois une adhérence forte et des pertes d'énergie faibles. On ne considérera que le cas où les roues sont en acier (comme le rail). Dans ce cas-là l'adhérence est alors assez faible ce qui limite l'accélération et le freinage. En revanche, un faible coefficient de friction et une grande raideur de l'acier assurent une dissipation d'énergie faible. À titre de comparaison, dans le cas des roues dotées de pneumatiques, le coefficient de friction est quatre fois plus grand et la résistance au roulement vingt fois plus grande que dans le cas de roues en acier.

Formulaire 1.1 - Expression des forces physiques
Les frottements efficaces :

 En traction : $F_{eff} \leq \min(+P/v, m \cdot g \cdot fac.)$

 En freinage : $F_{eff} \geq -m \cdot g \cdot fac.$
Les frottements inefficaces :

$$F_{ineff} = F_{ar} + F_{courbe} + F_{perm} + F_{défaut} + F_{air}$$

$$\left. \begin{array}{l} F_{ar} = -10^{-2} \cdot m \cdot g \quad \text{si } v = 0 \\ F_{courbe} = -m \cdot g \cdot e/r_c \\ F_{perm} = -10^{-3} \cdot m \cdot g \\ F_{défaut} = -10^{-4} \cdot v \cdot m \cdot g \\ F_{air} = -3 \cdot 10^{-6} \cdot v^2 \cdot m \cdot g \end{array} \right\} \text{en [N]}$$

La force de déclivité :

$$P_x = -m \cdot g \cdot \sin(\alpha)$$

Variables utilisées :

 P $\approx 3000 \text{ kW}$ puissance de traction

 fac. ≈ 0.2 coef. d'adhérence acier-acier

 m $\approx 2.10^5 \text{ kg}$ masse du train

 g $\approx 10 \text{ m.s}^{-2}$ accélération de la gravité

 e $= 1.4 \text{ m}$ écartement des voies

 r_c $> 150 \text{ m}$ rayon de courbure de la voie

 v $< 30 \text{ m.s}^{-1}$ vitesse du train

 $\alpha < 4^\circ$ inclinaison

À partir du formulaire 1.1, on est en mesure de tracer les graphes représentant les valeurs que peuvent prendre F_{eff} , F_{ineff} et P_x en fonction de la vitesse et de l'inclinaison. Ces graphes sont présentés dans la figure 1.2.

F_{eff} peut évoluer entre ses deux limites, maximum de traction et minimum de freinage. En pratique, on ne s'approche pas trop des limites d'adhérence (partie constante de ces courbes) et on garde une bonne marge de sécurité pour éviter les phénomènes de patinage et d'enrayage. En effet, le coefficient de frottement acier-acier peut varier par exemple selon les conditions météorologiques ou l'état des surfaces en contact.

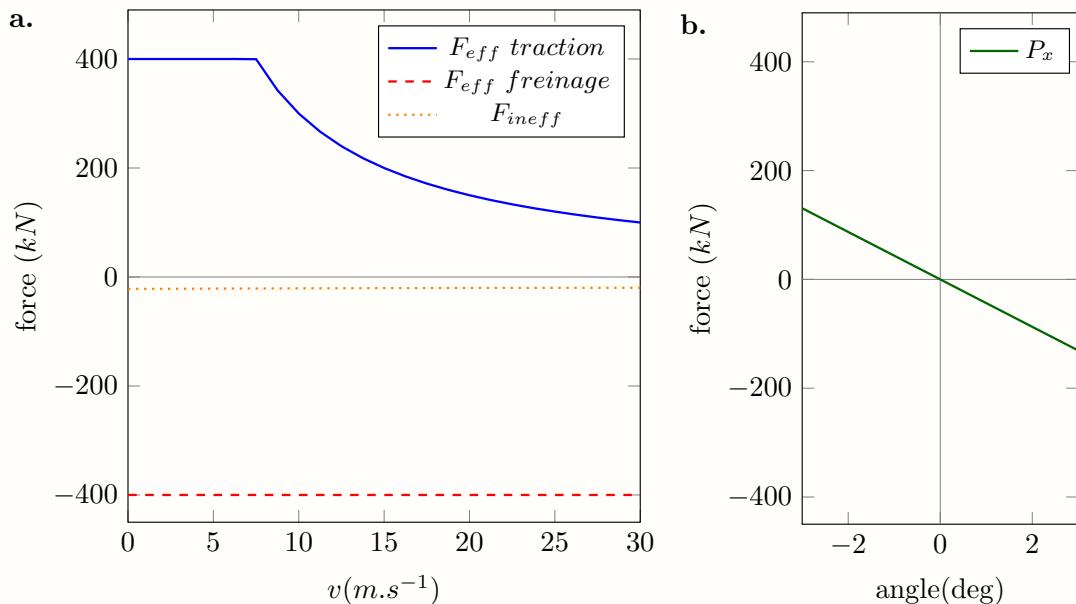
Pour tracer F_{ineff} on n'a pas tenu compte de F_{ar} qui est nul si la vitesse n'est pas nulle et qui influence donc peu la dynamique des trains. Concernant F_{courbe} on a tenu compte du fait que les courbes à petit rayon de courbure peuvent seulement être parcourues à faible vitesse. À titre d'illustration on a réglé la dépendance entre la vitesse et le rayon de courbure en prenant : $r_c = 150 \text{ m}$ pour $v = 0 \text{ m.s}^{-1}$ et $r_c = 300 \text{ m}$ pour $v = 30 \text{ m.s}^{-1}$. Ces valeurs sont les limites minimum que peut prendre le rayon de courbure à ces vitesses pour des raisons de sécurité³. Par conséquent, la force F_{ineff} que l'on a représentée est seulement une possibilité parmi d'autres. Ce choix correspond à ce que F_{ineff} pourrait valoir au maximum en valeur absolue. Ici la diminution de F_{courbe} au cours du temps l'emporte sur l'augmentation d'autres forces et F_{ineff} tend donc à diminuer avec la vitesse⁴.

3. Dans le terrible accident ayant eu lieu en Espagne en juillet 2013 à Saint-Jacques de Compostelle, une vitesse trop élevée par rapport au rayon de courbure de la voie était en cause.

4. Soulignons que c'est vrai uniquement parce qu'on considère un fort rayon de courbure à faible vitesse. En l'absence de virage, F_{ineff} augmente bien avec la vitesse.

On en conclut que F_{ineff} reste faible par rapport aux valeurs que peut prendre F_{eff} ⁵. Comme on vient de le dire, c'est l'un des avantages du ferroviaire sur la voiture, et cet avantage vient à la fois de la nature du contact acier-acier mais aussi de la forme effilée du train. Enfin on voit que la force de déclivité P_x peut vite devenir importante, elle est ici représentée pour des inclinaisons comprises entre ± 3 degrés, c'est-à-dire des déclivités comprises entre $\pm 0,5\%$. 0,5% de déclivité représente déjà une pente assez élevée en ferroviaire. La figure 1.2 sera utile pour comprendre la construction du modèle présenté au chapitre 2.

Figure 1.2 - Graphe des composantes de la force appliquée au train



Cette figure, basée sur le formulaire 1.1 représente les graphes de l'expression des forces F_{eff} , F_{ineff} et P_x en fonction de la vitesse et de l'inclinaison. a. F_{eff} peut évoluer entre ses deux limites maximum de traction et minimum de freinage. F_{ineff} est un exemple de ce que pourrait valoir cette force au maximum (en valeur absolue). b. La force de déclivité ne dépend pas de la vitesse et est représentée en fonction de l'inclinaison.

En pratique, la faible adhérence implique également que les distances de freinage et donc de sécurité soient plus grandes entre deux trains. Dans ces conditions, le freinage s'il est déclenché à la vue d'un train arrêté (ou de tout autre obstacle) ne garantit pas la sécurité et la collision devient inévitable au-dessus d'une certaine vitesse. Il faut donc utiliser un moyen qui permette aux conducteurs de prolonger leur vision au-delà de leur champ de vision. C'est le rôle du cantonnement et de la signalisation présentée ci-après.

5. La force aérodynamique peut cependant augmenter d'un ordre de grandeur dans les tunnels à cause de l'effet piston. Cela peut conduire F_{ineff} à ne plus être tout à fait négligeable pour des vitesses autour de $30 \text{ m.s}^{-1} \approx 110 \text{ km.h}^{-1}$. Par souci de simplicité on n'en tient pas compte bien que les tunnels en zone urbaine soient monnaie courante.

1.1.2 L'infrastructure

On a présenté dans la définition du chemin de fer l'infrastructure minimale que comporte un chemin de fer, à savoir la voie ferrée et un ensemble de points d'arrêt. Cette définition est volontairement minimale en ce qu'elle doit englober les nombreuses configurations existantes. Cette sous-section est dédiée à une description plus en profondeur de l'infrastructure ferroviaire. Nous complèterons la description de l'infrastructure avec le cantonnement, la signalisation, les contrôles de vitesses, les gares et les aiguillages. Nous discuterons également de l'aspect guidé du chemin de fer. Cette description apporte des éléments de compréhension nécessaires pour suivre la suite de la thèse et n'a pas vocation à être exhaustive. À plus grande échelle, on s'intéressera à l'organisation spatiale du réseau ferré.

Le cantonnement et la signalisation

Comme on vient de le voir, le cantonnement et la signalisation sont un moyen pour les conducteurs de train de recevoir des informations sur ce qui se passe bien au-delà de leur champ de vision. Cela les autorise à rouler à des vitesses élevées tout en leur assurant d'être prévenus suffisamment à l'avance de la présence d'un obstacle afin de réussir à arrêter leur train.

Le principe du cantonnement est de diviser une voie en segments appelés cantons et d'interdire à deux trains de se trouver sur le même canton. La signalisation permet de concrétiser cette interdiction en relayant l'information d'occupation d'un canton, ou de la présence d'un autre obstacle, sur plusieurs cantons. L'affichage de cette information est visible par le conducteur sous la forme d'un signal.

On distingue les types de cantonnement selon :

- Que les cantons soient fixes : avec une signalisation implantée au sol ou plus rarement en cabine pour augmenter la visibilité. C'est de loin le système le plus répandu.
- Que les cantons soient mobiles : ce qui implique une signalisation en cabine. C'est un système adapté aux grandes vitesses où des signaux implantés sur la voie seraient visible un laps de temps trop court pour le conducteur.
- Le nombre d'aspects disponibles pour les signaux. Plus il y en a, plus la distance d'arrêt entre un train et un obstacle est divisible. À mesure que l'on se rapproche d'un obstacle, les aspects correspondent à des vitesses de consigne de plus en plus faibles. Ces notions seront développées un peu plus tard dans ce chapitre.

Dans cette thèse on se focalisera sur un cantonnement fixe et une signalisation à trois aspects illustrée par la figure 1.3. Les trois aspects présentés sont les suivants :

- Le signal voie libre ou feu vert. Un train qui franchit ce signal à l'entrée d'un canton peut circuler à la vitesse maximum autorisée sur l'ensemble de ce canton. Ce canton et le suivant sont libres.
- Le signal d'avertissement ou feu jaune. Un train qui croise ce signal à l'entrée d'un canton doit être en mesure de s'arrêter avant le prochain signal s'il reste fermé. En pratique, le train doit ralentir jusqu'à atteindre une vitesse de sécurité, la VISA pour

Vitesse Sécuritaire d'Approche (très souvent 30 km.h^{-1}). Il ne la dépassera pas jusqu'au franchissement du signal suivant. Elle assure que l'arrêt du train soit possible avant un obstacle ou un signal fermé en freinant à vue. En pratique, on commence à freiner avant de franchir ce signal. Ce canton est vide mais le suivant est occupé.

- Le signal fermé dont il existe 2 variantes : le sémaphore impose l'arrêt du train mais autorise ensuite le franchissement à très faible vitesse, il signifie que le canton est occupé. Le carré impose l'arrêt et interdit le franchissement, ce signal protège un endroit particulier comme un aiguillage ou des travaux. En pratique, les sémaphores sont rarement franchis. On parlera de feu rouge.

Le contrôle de vitesse

Les contrôles de vitesse sont des équipements de sécurité situés à bord des trains et qui servent à contrôler la vitesse du train. Ces équipements qui ont des taux de défaillance bien élevés à ceux d'un humain servent à vérifier que le conducteur respecte les recommandations de vitesse. En pratique, dans la configuration de la figure 1.3.c, si le conducteur tarde à freiner, ou ne freine pas assez, le contrôle de vitesse va l'avertir par une alarme, puis, s'il ne réagit toujours pas, déclencher un freinage d'urgence.

Nous présentons deux types de contrôle de vitesse mais il faut savoir que d'autres variantes existent :

- Le KVB : Contrôle de Vitesses par Balises. Au franchissement d'un signal, une balise communique avec le train pour informer le KVB de l'aspect du signal franchi, ce dernier est alors en mesure de calculer la courbe de la vitesse limite à ne pas dépasser. Si c'est un signal jaune, le conducteur est tenu de ralentir à la vitesse de sécurité et de ne plus la dépasser jusqu'au franchissement du signal suivant même si ce signal passe au jaune puis au vert entre temps.
- Le KVBP : Contrôle de Vitesse par Balises pour les Prolongements. La transmission entre le train et l'information sur l'état des signaux est permanente et ne se limite pas à l'endroit de l'emplacement des balises. Le KVBP peut relâcher les contraintes de vitesse si le danger de collision s'est éloigné. En pratique, le système n'informe pas le conducteur qui ne pourra s'affranchir de la vitesse de sécurité qu'au moment où il aperçoit le signal suivant au jaune ou au vert. Cette précaution augmente le niveau de sécurité en assurant que le KVBP et le conducteur vérifie tous les deux que le signal suivant n'est plus rouge. Cela contribue à augmenter la sécurité en palliant la défaillance éventuelle d'un des deux systèmes.

Les gares

Les gares sont des points d'arrêt possibles pour les trains. Leur fonction est de permettre aux voyageurs de monter et de descendre du train. C'est une zone d'échange de voyageurs entre un train et un autre train⁶ ou entre un train et le monde extérieur.

6. On parle alors de correspondance.

Les aiguillages

Un aiguillage est un système mécanique qui permet à un train de changer de voie. Il faut noter que lorsqu'un train se présente devant un aiguillage, la décision de la direction à prendre ne lui appartient pas, ce qui représente une différence notable avec ce qui se passe en trafic routier lorsqu'un choix de direction se présente. Pour des raisons de sécurité, la signalisation interdit l'accès à un aiguillage dont l'aiguille est en mouvement.

Le rail et le guidage

La nature guidée du chemin de fer résulte des points de contact entre la roue et le rail, la trajectoire est imposée, on dit qu'il y a un seul degré de liberté. Le chemin de fer s'oppose sur ce point au transport routier et au transport aérien où il y a respectivement deux et trois degrés de liberté dans le cas général.

Le concept de degré de liberté d'un véhicule est relié au nombre de variables nécessaire pour décrire son mouvement. Précisons ce concept de degré de liberté avec des exemples issus de trois modes de transport différents :

- La trajectoire d'un avion est décrite avec trois variables, par exemple : l'altitude, la latitude et la longitude. Pour un avion, évoluer dans cet espace à trois dimensions permet d'adopter un grand nombre de trajectoires pour contourner un obstacle ; ainsi une zone d'orage, de conflit militaire ou la trajectoire d'un autre avion peuvent facilement être évitées. Cependant, les phases de décollage et d'atterrissement dans les aéroports impliquent des contraintes spatiales fortes qui réduisent localement les trajectoires à un seul degré de liberté. Les aéroports sont des zones où l'on observe une interaction entre les avions qui se manifeste dans des phénomènes de files d'attente.
- La trajectoire d'une voiture est décrite avec deux variables, une première coordonnée donne le point kilométrique de la route où la voiture se situe, une seconde donne la position de cette voiture sur la largeur de la route. Ce second degré de liberté obéit à une contrainte forte : elle est limitée par la largeur de la route à quelques mètres. En pratique, cette contrainte peut rendre le dépassement impossible comme dans le cas d'une rue étroite, on considère alors qu'il n'y a qu'un seul degré de liberté. Au contraire, un grand boulevard, une voie rapide ou une autoroute sera suffisamment large pour autoriser à tout moment le dépassement.
- La trajectoire d'un train est décrite avec une seule variable : le point kilométrique de la voie sur lequel il se situe. Cette configuration interdit tout dépassement sur une voie de chemin de fer.

Pour un véhicule, il faut au moins deux degrés de liberté dans la trajectoire pour permettre un dépassement. Toutefois il peut être rendu difficile voire impossible lorsque les contraintes sur le second degré de liberté sont fortes comme c'est le cas dans certaines configurations observées en trafic routier. Cette capacité à dépasser est importante car

c'est un facteur de résilience⁷ pour le trafic : la possibilité de contourner un obstacle permet de réduire voire d'annuler son impact sur le trafic. De ce point de vue, le trafic ferroviaire est moins résilient que le trafic aérien ou routier.

Un autre caractère de résilience du trafic réside dans la façon dont sont agencées les différentes voies entre elles, qu'elles soient routières ou ferrées. On parle d'une organisation en réseau.

Organisation spatiale d'un réseau ferré

L'organisation spatiale d'un réseau de transport raconte une histoire sur la dynamique des déplacements et sur les enjeux économiques, politiques et financiers des territoires desservis. À l'instant présent, un réseau est modelé par l'héritage du passé sur lequel il s'appuie pour faire face à une demande de transport. Un réseau de transport n'est pas seulement destiné à satisfaire la demande des passagers, il influe aussi sur l'évolution de la demande future et facilite le développement économique des territoires les mieux desservis. On dit à ce titre qu'un réseau de transport est un outil d'aménagement du territoire, les enjeux économiques, politiques et financiers du présent façonnent donc à leur tour le futur du réseau de transport. Dans le cadre de cette thèse, nous garderons à l'esprit que les problématiques ferroviaires s'inscrivent dans un contexte historique, mais nous nous intéresserons principalement au fonctionnement du réseau de transport ferré à l'instant présent.

Si du point de vue de la possibilité de dépasser, c'est-à-dire à l'échelle de la taille d'un véhicule on considère que la trajectoire des voitures s'effectue avec deux degrés de liberté, à l'échelle du réseau, on ne distingue plus les quelques mètres de largeur des voies et on peut les considérer comme des objets à une seule dimension. La finalité d'un réseau de transport étant de couvrir un territoire et donc une surface, la question qui se pose est la suivante : comment desservir au mieux l'espace (à deux dimensions) avec un ensemble de voies qui sont des objets à une dimension ?

On va maintenant définir de manière succincte ce qu'est la topologie d'un réseau de transport. Après avoir présenté quelques types usuels de topologie, on discutera de leur façon de couvrir l'espace. Nous verrons aussi comment la topologie d'un réseau de transport influe sur sa résilience.

7. Comme nous n'introduisons pas ici de mesure de cette résilience, nous prenons une définition simple et générale : *la résilience désigne la capacité pour un corps, un organisme, une organisation ou un système quelconque à retrouver ses propriétés initiales après une altération.* ([HOLLNAGEL et collab. \[2009\]](#); [TÖYLI et collab. \[2013\]](#)).

Définition - Topologie d'un réseau de transport

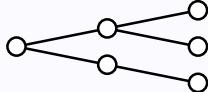
Un réseau de transport est souvent représenté par des arcs ou segments qui représentent les voies de circulation et qui relient entre eux des nœuds qui représentent des bifurcations ou des gares.

La topologie d'un réseau décrit la façon dont les nœuds sont reliés entre eux par ces segments sans prendre en considération les distances réelles. On distingue trois topologies types utiles pour la suite :

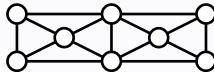
- **Le type linéaire**



- **Le type en arbre ou ramifié**



- **Le type maillé**



- Le type linéaire offre une couverture spatiale limitée, la défaillance d'un seul nœud coupe le réseau en deux sous-réseaux indépendants.
- Le type en arbre offre une meilleure couverture spatiale, il est souvent utilisé pour une ligne de chemin de fer qui, partant d'un centre urbain, se ramifie à mesure qu'elle s'en éloigne. Ici encore la défaillance d'un seul nœud coupe le réseau en deux sous-réseaux indépendants.
- Le type maillé est bien plus adapté pour couvrir l'espace. Il offre de plus la possibilité de contourner un nœud défaillant ce qui est un facteur de résilience important. En revanche, le nombre de segments par nœud augmente, ce qui implique la construction de plus de voies et donc un coût de construction et de maintenance plus élevé.

Les réseaux de transport ne relèvent pas forcément d'un seul type de topologie mais souvent d'une combinaison de plusieurs types. La figure 1.4 présente une comparaison sur un exemple entre la topologie d'un réseau urbain routier et un réseau urbain ferré.

Le réseau routier présenté à la figure 1.4.a est un réseau maillé. Il offre la possibilité de contourner un accident, des travaux ou une zone de ralentissement qui paralyserait un nœud ou un segment de ce réseau. Ainsi, même quand il y a une impossibilité de dépasser un autre véhicule sur un même segment, le maillage offre des possibilités de le faire en changeant son itinéraire. Le caractère résilient qui découle de cette possibilité reste toutefois à quantifier : ce n'est pas parce que le réseau offre la possibilité de contourner un nœud ou un segment défaillant que ce contournement ne se fait pas avec une dégradation relative du confort et du temps de parcours pour les passagers.

Pour un réseau de transport ferré, la vision n'est pas la même selon que l'on regarde se déplacer les passagers ou les trains. Du point de vue des passagers, le réseau ferré présenté à la figure 1.4.b contient une seule maille et plusieurs ramifications. La topologie de ce réseau apparaît également comme linéaire à certains endroits comme au sud de la ligne orange. Cette vision est valable car les passagers sont capables de changer de train dans les gares offrant une correspondance. Si en revanche on regarde le réseau du point de vue des trains, il faut considérer les trois lignes comme autant de réseaux indépendants car un train de la ligne bleue ne peut pas emprunter les voies des autres lignes pour

contourner une coupure en son sein comme pourrait le laisser croire la topologie du réseau. La structure topologique de ces lignes est linéaire pour les lignes orange et bleues et à la fois linéaire et ramifiée pour la ligne rouge. Il y a donc aux endroits où le maillage le permet, une certaine résilience du point de vue des passagers qui n'existe pas du point de vue des trains⁸. Encore une fois la possibilité pour les passagers de contourner un obstacle entraîne à priori un rallongement du temps de transport et souvent des conditions de transports dégradées et le caractère résilient qu'offre la topologie du réseau n'est que relatif.

À l'échelle locale, il arrive qu'il y ait un dédoublement des voies. La figure 1.4.c présente un tel dédoublement des voies au niveau des quais d'une gare comme c'est fréquemment le cas. Cette configuration constitue une maille du point de vue topologique. Elle permet donc en utilisant la voie qui passe par le second quai de dépasser un train qui serait bloqué ou qui devrait patienter sur le premier quai. Cette maille offre donc un peu de résilience à la ligne si un train est stoppé dans une telle gare.

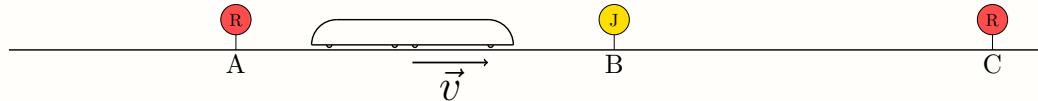
Pour revenir sur des cas plus concrets la figure 1.5 représente la carte du réseau régional de la région parisienne (qui fournit les données utilisées dans cette thèse). Cette carte n'est pas à l'échelle. Ce réseau est partagé entre deux opérateurs : la RATP (Régie Autonome des Transports Parisiens) gère une partie de la ligne B (en bleu) et une partie de la ligne A (en rouge). Transilien SNCF gère toutes les autres lignes.

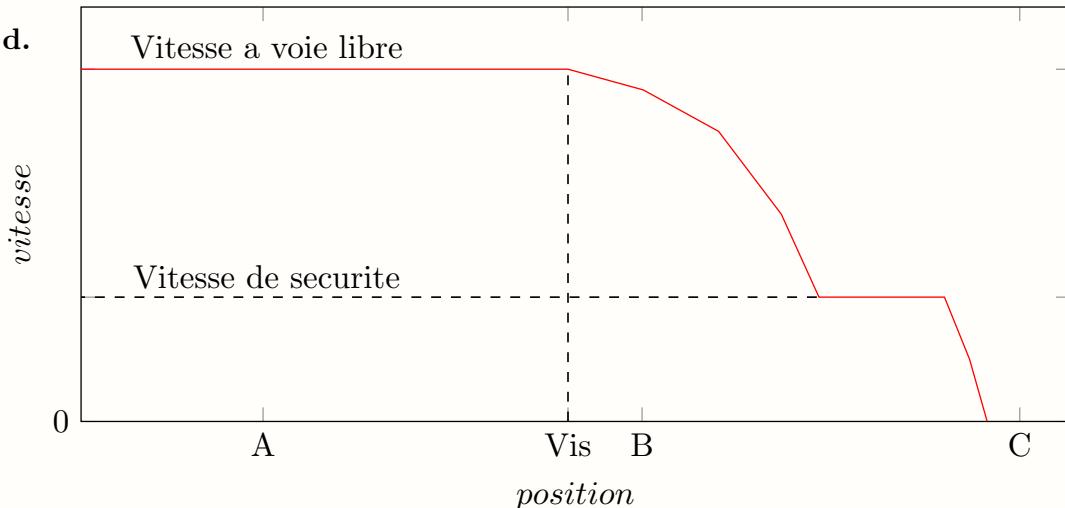
Contrairement au réseau du métro parisien qui présente un maillage assez fort, le réseau régional possède des lignes avec une topologie de type linéaire qui a tendance à se ramifier lorsque l'on s'éloigne du centre de Paris. Ce réseau régional comporte toutefois quelques boucles. Il faut savoir que la forte interconnexion entre ces deux réseaux conduit à un système de transport de très bonne qualité dans Paris, qualité que l'on ne retrouve pas à mesure que l'on s'éloigne de Paris. En effet, la densité de gares diminue, le maillage se dégrade entraînant une perte de résilience et l'offre de transport devient essentiellement radiale depuis le centre de Paris vers la banlieue, l'offre sur les déplacements périphériques de banlieue à banlieue reste alors essentiellement routière.

On vient de voir que le caractère résilient d'un réseau de transport est influencé par la possibilité de dépasser, mais aussi par l'organisation topologique de ce réseau. Un réseau de transport dont la résilience serait plus faible pourrait offrir une même qualité de service à ses voyageurs si par ailleurs il diminuait la fréquence des incidents qui affectent un nœud ou un segment en son sein. C'est dans un contexte d'amélioration de l'offre de transport d'un réseau ferré urbain que se situe cette thèse. Nous venons de comparer plusieurs fois le transport ferré et le transport routier, il faut garder à l'esprit que ces deux modes de transport s'opposent par bien d'autres aspects. Nous allons d'ailleurs présenter maintenant un aspect spécifique au trafic ferroviaire et non le moindre, la planification.

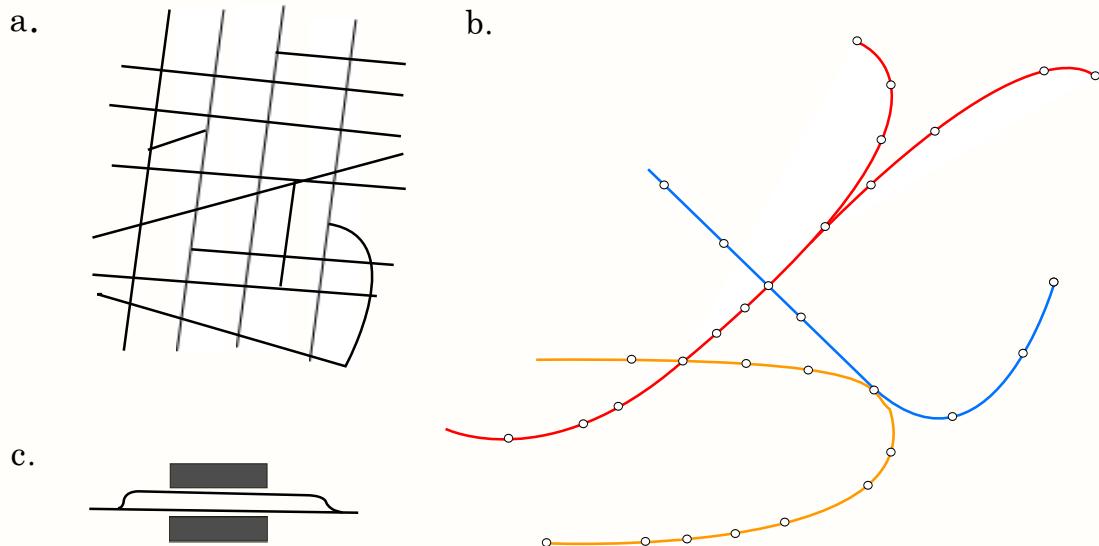
8. Le réseau tel qu'il est vu par les passagers est en réalité encore plus complexe et tient compte des possibilités de changer de mode de transport en cours de route. C'est une vision multimodale du transport.

Figure 1.3 - Exemple de signalisation à trois aspects
a.

b.

c.

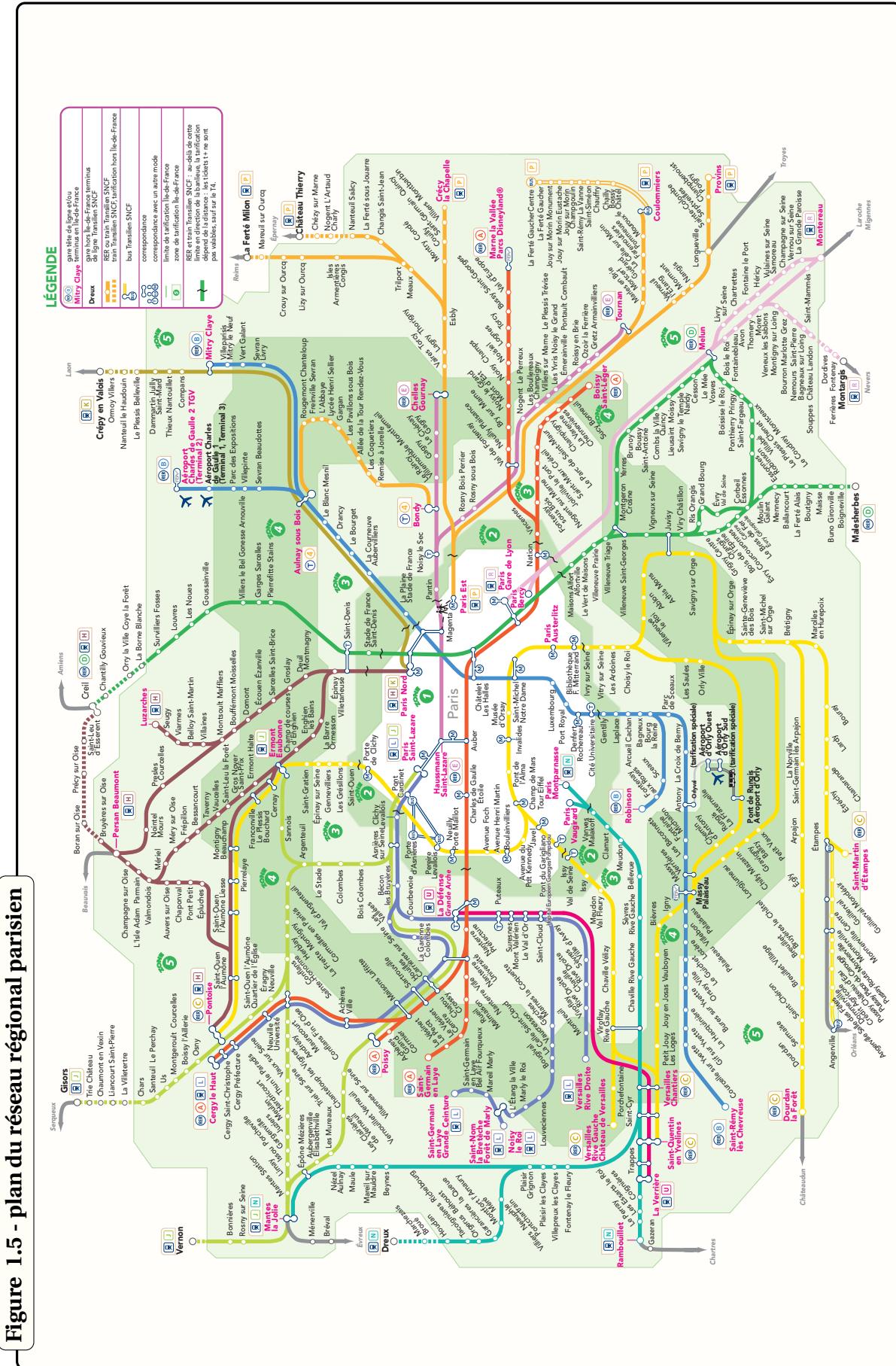
d.


Cette figure présente le fonctionnement de la signalisation trois aspects à cantonnement fixe. C'est un cas particulier dans la mesure où on a décidé que le signal C ne repasse pas au jaune (J) ou au vert (V). Le passage de la figure **a** à la figure **b** montre que le signal A passe au rouge (R) lorsqu'on le franchit, le passage de la figure **b** à la figure **c** montre que lorsque qu'un train quitte un canton, le signal A à l'entrée de ce canton passe du rouge au jaune. La figure **d** montre un exemple de ce que peut être la courbe de la vitesse en fonction de la position sur les cantons successifs. Cette courbe présente également les 3 vitesses de consignes correspondant aux trois signaux : la vitesse à voie libre, la vitesse de sécurité et la vitesse nulle. On remarquera que le freinage est amorcé au point Vis dit de visibilité du signal B et non à son franchissement.

Figure 1.4 - Comparaison de réseaux urbains routier et ferroviaire

Cette figure permet de comparer la topologie de deux réseaux urbains routier et ferroviaire. Les réseaux présentés, bien qu'ils soient inspirés de situations observées en région parisienne, sont fictifs. La figure **a** représente un réseau routier urbain. Ce réseau est très maillé. La figure **b** représente un réseau urbain ferroviaire à grande échelle, les couleurs correspondent à trois lignes différentes. Les gares sont représentées. La figure **c** représente un dédoublement local d'une voie ferrée dans une gare.

Figure 1.5 - plan du réseau régional parisien



1.2 Le chemin de fer : gestion des circulations

1.2.1 Planification : gestion à l'horaire ou à la fréquence

Définition - Planification

La planification ou la programmation est un procédé qui permet de définir et d'organiser à l'avance, les différentes tâches à réaliser, les ressources utilisées ainsi que l'ordre dans lequel il faut les réaliser, en vue d'atteindre un objectif.

L'objectif qui nous préoccupe ici est l'acheminement des passagers jusqu'à la destination désirée. Il est intéressant de noter que cette tâche ne passe pas par une planification globale des déplacements dans la gestion d'un réseau routier. Chaque usager du réseau de transport est libre de planifier ou non son déplacement. En revanche, en ferroviaire, la planification est forte. Pour ces derniers, on distingue deux modes de planification :

- Si l'on prévoit que le temps d'attente entre deux trains sera élevé ($> 15\text{ min}$), on choisit généralement de gérer les trains à l'horaire pour réduire les temps d'attente des usagers en gare. Cela demande une planification lourde et dont la révision demande beaucoup de travail.
- Si l'on prévoit que le temps d'attente entre deux trains sera faible ($< 10\text{ min}$), on choisit généralement de gérer les trains à la fréquence. La planification est alors plus légère. On parle alors souvent de métro et non de train.

Il faut noter que le temps d'attente entre deux trains n'est pas le seul paramètre qui influence le choix de gestion à l'horaire ou à la fréquence. Le trafic ferroviaire étant organisé en lignes, la topologie de ces lignes peut entraîner une difficulté à concevoir une gestion à la fréquence. En effet, à chaque fois qu'il y a une convergence entre deux branches (au sein d'une même ligne ramifiée ou lorsque deux lignes partagent la même portion de voie), la gestion à la fréquence ne permet pas d'éviter les conflits qui forceraient un train à attendre que l'autre passe. Seule la planification d'une grille horaire permet de synchroniser ces deux flux qui convergent et, de manière générale, de s'assurer que chaque train ne rencontre que des signaux ouverts à mesure qu'il avance.

Dans ce manuscrit de thèse nous nous sommes intéressés à un système ferroviaire avec une planification à l'horaire uniquement. Un système planifié à l'horaire peut en outre être cadencé. Il s'agit de répéter un motif horaire simple tout au long de la journée sur l'ensemble ou sur une partie du réseau. Cette disposition facilite grandement la lecture des horaires par les passagers et synchronise les correspondances entre les trains sur le réseau sur lequel il est appliqué. En revanche, le cadencement représente une contrainte supplémentaire lors de la construction des horaires. Dans ce manuscrit nous ne nous intéressons ni au point de vue des passagers ni aux correspondances, ce qui limite l'intérêt de parler du cadencement. Nous faisons donc le choix de ne plus parler du cadencement dans la suite,

ce qui ne veut pas dire que les situations étudiées ne pourront pas être cadencées. Nous allons maintenant présenter plus en détails comment fonctionne un système planifié.

1.2.2 Les sillons

Définition - Graphique de circulation et sillon

Le caractère mono-dimensionnel du chemin de fer nous autorise à représenter le déplacement des trains sur un graphique à deux dimensions, l'une spatiale et l'autre temporelle. Une telle représentation est appelée graphique de circulation.

Le couple espace-temps affecté par le gestionnaire d'infrastructure à la circulation d'un train est appelé un sillon.

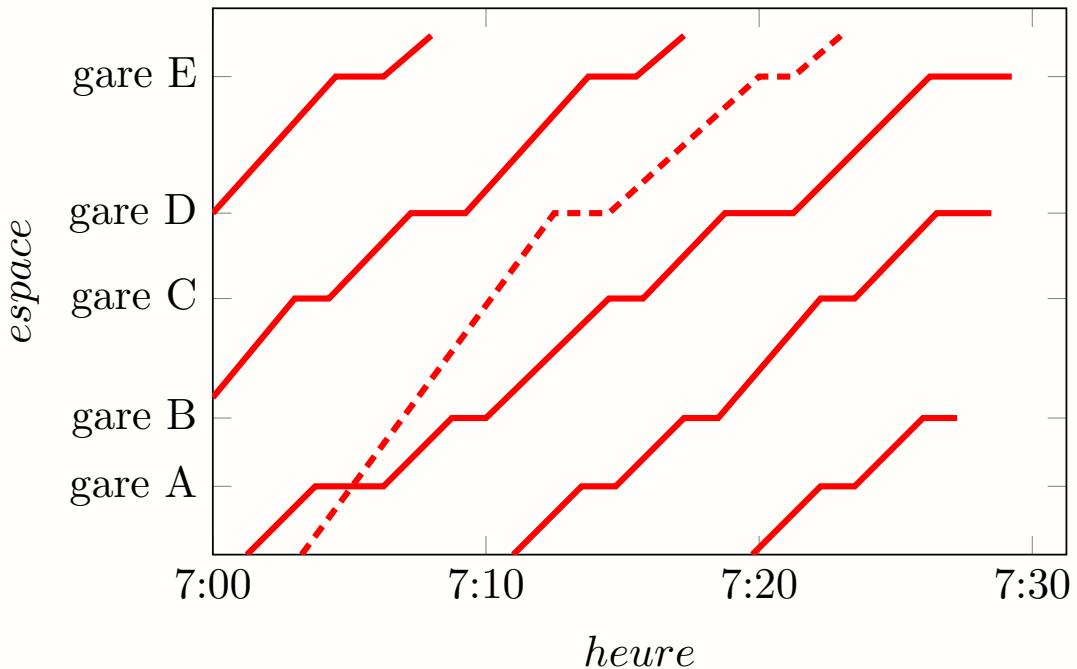
On distingue le graphique de circulation que l'on trace avec les sillons et qui constitue le plan de transport du graphique de circulation réel où l'on représente les déplacements réels des trains.

En réalité le plan de transport prend également en compte, le planning des conducteurs et du matériel roulant. Dans ce manuscrit nous n'abordons pas ces sujets. Nous assimilons donc le plan de transport au graphique de circulation des trains.

Un exemple fictif de graphique de circulation avec ses sillons correspondants est présenté sur la figure 1.6. Six sillons sont représentés, l'un d'entre eux correspond à un train qui ne marque pas tous les arrêts en gare et profite de la gare A pour dépasser un autre train. Comme vu précédemment, ce dépassement suppose que la gare peut faire passer le train sur un quai supplémentaire.

Bien qu'il ait été transposé du papier à l'écran d'ordinateur, cet outil a peu évolué depuis ses débuts. C'est un outil visuel, efficace, utilisé pour contrôler le processus de planification des horaires mais aussi pour permettre le suivi en temps réel des circulations et pour repérer les trains qui ne respectent pas l'horaire. Dans ce cas, on superpose sur le graphique de circulation les sillons des trains avec leur tracé en temps réel.

Il convient de remarquer que les sillons sont ici représentés par un trait d'épaisseur nulle. C'est une approximation voulue par l'échelle d'observation : on regarde plusieurs sillons en même temps. Si l'on regarde de plus près, le tracé du sillon fera apparaître son épaisseur (cf figure 1.8). En effet, le sillon représente l'occupation de l'espace temps qui tient compte de la taille du train ainsi que des cantons dont les signaux ne sont pas au vert car ils sont en protection de ce train. Nous verrons dans la section suivante comment la largeur des sillons tient compte de ces paramètres et comment cela influe sur les calculs de capacité et de débit.

Figure 1.6 - graphique de circulation

Ce graphique de circulation fictif a été inspiré de graphique réel en provenance du réseau Transilien. Il présente tous les sillons entre 7 :00 et 7 :30 sur un trajet comprenant cinq gares et pour un seul sens de circulation. À cette échelle la taille des trains est négligeable et ces sillons sont représentés par un simple trait. On remarque qu'un train, dont le sillon est représenté par un trait discontinu, en dépasse un autre vers 7:05 au niveau de la gare A.

1.2.3 Débit et capacité

Définition - Le débit

En physique, un débit mesure une quantité de matière qui traverse une surface donnée par unité de temps.

Par analogie, on définit le débit ferroviaire en un point du réseau comme un nombre de trains ou de passagers qui franchit ce point par unité de temps.

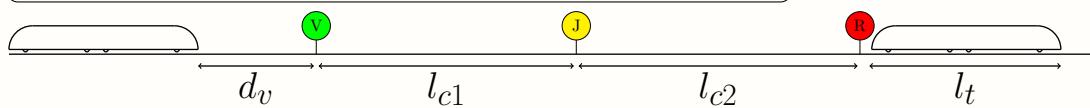
En pratique, le passage des trains étant un phénomène discontinu, on ne peut pas mesurer de débit instantané (i.e. sur une plage de temps très courte). On mesurera des débits moyens sur des périodes suffisamment longues pour pouvoir compter suffisamment de trains.

Définition - La capacité

On définit la capacité d'une portion du réseau comme le débit maximum possible à la traversée de cet élément indépendamment du débit qui serait imposé par les autres portions.

Cette définition dépend donc de l'échelle de temps que l'on choisit pour mesurer le débit, mais aussi de la taille de la portion de réseau considérée. Prendre une portion trop petite n'aurait pas de sens, car comme on va le voir il faut au moins deux cantons pour évaluer un débit maximum. Cette définition peut aussi dépendre du matériel roulant considéré et des missions réalisées.

Il faut savoir que d'autres définitions du débit et de la capacité sont parfois utilisées à la SNCF. Cependant, dans la suite de ce manuscrit nous tiendrons à ces définitions qui sont cohérentes avec celles utilisées en physique ou dans d'autres domaines des transports comme le trafic routier. À ce propos, notre choix de définitions est destiné à faciliter la compréhension des parallèles faits entre le trafic ferroviaire et le trafic routier dans la suite de ce manuscrit. Notons que cette définition de la capacité d'une portion du réseau n'est pas incompatible avec le fait que cette capacité ne sera jamais atteinte si une autre portion du réseau n'autorise qu'une capacité plus faible.

Figure 1.7 - distance minimale prévue entre deux trains


Ce schéma présente la distance minimale qu'il faut garder entre deux trains pour assurer de continuer de rouler à voie libre. L_{c1} et L_{c2} sont les longueurs des deux cantons, l_t est la longueur d'un train et d_v la distance de visibilité du signal.

La figure 1.7 présente un schéma permettant de comprendre le calcul de la capacité d'une portion de ligne. Considérons cette portion constituée d'une suite de quelques cantons. Soit v la vitesse à laquelle circulent les trains à voie libre sur cette portion et $v_{\text{sécu}}$ la vitesse sécuritaire d'approche telle que définie plus tôt. Pour pouvoir rouler à la vitesse v , la signalisation impose entre les trains une distance minimum d_{\min} :

$$d_{\min} = d_v + l_{c1} + l_{c2} + l_t$$

Avec l_{c1} et l_{c2} la taille de deux cantons voisins, l_t la taille d'un train et d_v la distance de visibilité du signal. Dans cette situation (ou la vitesse de circulation est homogène) le débit maximal est donné par l'expression (1.1) :

$$D_{\max} = v / d_{\min} = v / (d_v + l_{c1} + l_{c2} + l_t) \quad (1.1)$$

Avant de conclure en toute rigueur il faudrait évaluer le débit maximal donné par la situation où les trains roulent tous à la vitesse de sécurité v_{secu} donc moins vite mais avec une distance moins grande entre les trains. En pratique, cette situation conduit à un débit maximal inférieur au débit maximal calculé ci-dessus et qui correspond donc à la définition de la capacité.

On a fait l'hypothèse que v était constante sur la portion de voie considérée. Ce n'est évidemment pas toujours le cas et la formule (1.1) est un cas particulier. La formule du débit maximal dans le cas d'une vitesse variable est simple à établir, il faut prendre en compte les temps de parcours $t_p(_)$ de chacune de ces quatre distances, c'est ce qui est fait dans la figure (1.2) :

$$t_p = t_p(d_v) + t_p(l_{c1}) + t_p(l_{c2}) + t_p(l_t)$$

$$D_{max} = 1/t_p = 1/(t_p(d_v) + t_p(l_{c1}) + t_p(l_{c2}) + t_p(l_t)) \quad (1.2)$$

Donc, à vitesse constante, sur une portion de ligne qui comporte plus de deux cantons, le calcul du débit est limité par le plus grand temps de parcours parmi toutes les paires de cantons consécutifs. Si la vitesse n'est pas constante, le débit sera limité par la paire de cantons (consécutifs) qui a le temps de parcours le plus grand. Dans le cas où cette portion de ligne comporte une gare, il faudra tenir compte non seulement des phases de freinage et d'accélération mais aussi du temps de stationnement du train à quai. Il en découle que la capacité d'une grande portion de ligne est le minimum des capacités que l'on peut calculer sur des subdivisions de cette portion. Les cantons sont dimensionnés de manière à optimiser le débit lorsque c'est possible, une vitesse de circulation réduite à voie libre implique des cantons plus courts.

La définition de la distance de visibilité du signal d_v fait l'objet de controverses à la SNCF. Nous considérons implicitement dans mon calcul de la capacité que c'est la distance minimale raisonnable à laquelle un train qui roule à voie libre anticipe le signal qu'il aperçoit devant lui. En pratique, lors de la conception des horaires, la distance effectivement prise en compte repose sur un temps forfaitaire : le temps minimal de visibilité de l'indication à voie libre, noté Khi . L'application de ce temps forfaitaire conduit à concevoir la ligne en dessous de ce que nous appelons la capacité. Nous allons justement maintenant voir quelques aspects de la construction du graphique de circulation théorique aussi appelé plan de transport.

1.2.4 Construction du plan de transport et robustesse

La construction d'un plan de transport est un processus complexe qui inclut de nombreuses étapes. La fonction socio-économique du chemin de fer est dans le cas qui nous préoccupe, le transport des passagers. La première étape de la construction de ce plan consiste à connaître la demande pour pouvoir construire une offre de transport adaptée. Une fois la demande évaluée, la construction du plan de transport doit tenir compte de nombreuses contraintes parmi lesquelles la disponibilité du matériel roulant et des conducteurs : il ne suffit pas d'avoir un sillon disponible, il faut également s'assurer que

le train et le conducteur sont disponibles au bon endroit au bon moment pour assurer le service correspondant. Mais laissons les contraintes logistiques de côté pour nous concentrer sur le tracé des sillons.

Le graphique de circulation

La figure 1.8 présente le détail d'un graphique de circulation où trois sillons sont visibles. On distingue sur le tracé de ces sillons un double trait, la distance verticale entre ces deux traits correspond à la longueur d'un train. L'axe des ordonnées indique également la position des signaux de cantonnement. L'occupation des cantons dans l'espace-temps est signalé par des rectangles rouges, de la même façon des rectangles jaunes signalent les cantons qui précèdent les cantons occupés et dont l'entrée est protégée par un signal jaune. Après les rectangles jaunes vient une troisième zone plus petite et qui n'est pas représentée pour ne pas surcharger la figure. Elle correspond au temps de visibilité du signal évoqué précédemment. En réalité, bien que ce temps de visibilité soit un élément important de dimensionnement du trafic, on pourra le négliger dans la partie théorique de ce manuscrit. En effet, quand on étudie les effets du cantonnement sur le trafic ferroviaire, le temps de visibilité du signal n'apparaît que comme un raffinement qui diminue le débit maximum mais ne change pas les mécanismes d'apparition de la congestion. Enfin la zone encadrée en pointillés et grisée correspond à la situation présentée dans la figure 1.7. Il s'agit d'une photographie du système pris à un instant t .

Construire le plan de transport consiste à trouver un graphique de circulation sur l'ensemble de la ligne pour lequel on respecte les contraintes de disponibilité du matériel et du personnel. De plus ce plan doit offrir un service de transport au plus proche des attentes des voyageurs. Ce plan de transport est tracé à voie libre. Cela signifie que l'on doit toujours laisser sur la figure 1.8 une marge temporelle entre un rectangle jaune et le rectangle rouge du train suivant. Cette marge correspond au temps de visibilité du signal vu précédemment. La capacité de la ligne est atteinte si les trains se succèdent au plus près les uns des autres sur le graphique tout en respectant cette marge. La figure 1.8 nous montre par ailleurs que le canton situé entre les points C_1 et C_2 limite cette capacité. Ce graphique de circulation n'est pas complet il est possible qu'une autre limitation plus forte de la capacité soit présente ailleurs sur la ligne. Le travail de construction du graphique ne peut se faire que de manière globale, ce qui rend la tâche ardue. Nous allons maintenant voir qu'il peut être intéressant de prendre en compte des marges supplémentaires entre deux trains au détriment du débit de la ligne.

La robustesse

Un train peut être perturbé par un certain nombre d'événements qui l'amènent à prendre du retard⁹. Dans ces conditions, il convient de définir des marges additionnelles. Les concepts de stabilité, de robustesse et de performance deviennent importants pour comprendre ce qui motive l'ajout de ces marges. Concernant ces trois concepts, il existe beaucoup de définitions. Nous allons tenter de les définir dans ce chapitre uniquement de la

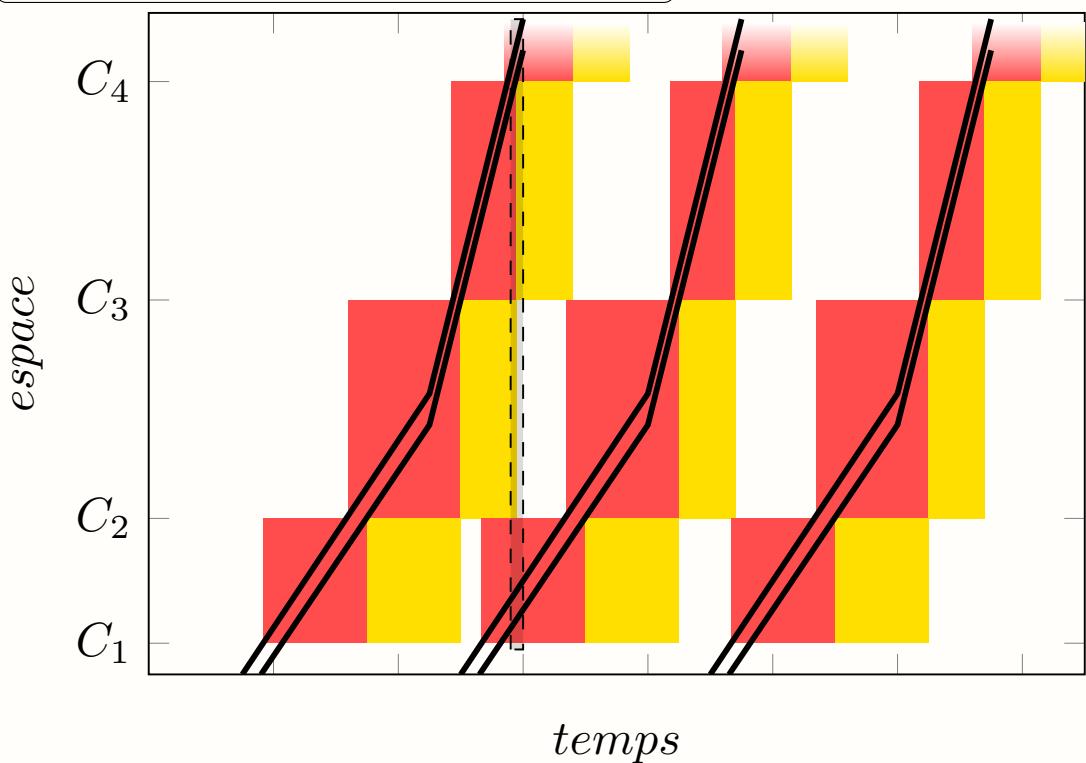
9. ou de l'avance qui, contrairement au retard sera assez faible mais peut rester, on le verra plus tard, gênante.

manière la plus générale possible. De cette façon ces définitions resteront compatibles avec les autres définitions en usage. Nous verrons dans le chapitre 2 que la quantification de ces concepts nous sera nécessaire ce qui impose de choisir une définition plus précise et d'abandonner ce consensus.

Définition - La performance

On dit d'un système qu'il est performant s'il remplit les objectifs pour lesquels il a été conçu.

Figure 1.8 - détail d'un graphique de circulation



Cet agrandissement d'un détail d'un graphique de circulation permet de rendre compte des effets de la taille du train et de la position des cantons. Les trois sillons sont ici représentés par des traits doubles noirs. La distance verticale entre ces deux traits correspond à la taille d'un train. Les positions C_1 à C_4 représentent les positions des signaux et le passage d'un canton à un autre. Les zones rouges montrent les cantons occupés au cours du temps, les zones jaunes montrent que le canton suivant est protégé par un signal jaune sur la même période. Tout le reste du graphique correspond à des cantons dont les signaux sont ouverts. La zone encadrée en pointillés corresponds à la figure 1.7

Définition - La stabilité

On dit d'un état d'équilibre d'un système dynamique qu'il est stable si, après de petites perturbations, ce système reste au voisinage de cet état d'équilibre.

Définition - La robustesse

La robustesse est définie comme la stabilité de la performance d'un système. L'état stable est alors l'état qui regroupe toutes les situations d'un système dans lesquelles il peut être considéré comme performant. On peut quantifier la robustesse par la mesure de l'amplitude minimum des perturbations nécessaires pour qu'un tel système ne soit plus performant.

On comprend bien que l'on peut obtenir des définitions plus précises dès lors que l'on précise l'objectif de performance de notre système, ainsi que la mesure des perturbations de ce système.

En ce qui concerne notre système, l'objectif est de réaliser le plan de transport en autorisant chaque train à faire des écarts à l'horaire tant que ceux-ci n'impactent aucun autre train et que ces écarts restent faibles. Remplir la grille horaire avec le maximum de trains de manière à ce que le débit atteigne la capacité n'engendre pas un plan de transport robuste. Pour pallier ce problème, des marges supplémentaires sont ajoutées dans la grille horaire. On distingue deux stratégies principales utilisées pour répartir ces marges :

- Soit de petites marges sont appliquées de manière équitable pour toutes les circulations.
- Soit des marges plus grandes sont appliquées mais uniquement toutes les quelques circulations.

La construction du plan de transport s'inscrit donc dans la recherche d'un compromis entre la maximisation du débit de la grille horaire et sa robustesse. L'optimisation d'une de ces deux grandeurs se faisant de manière générale au détriment de l'autre. Pour guider ceux qui construisent les plans de transport, il existe des outils informatiques qui permettent de simuler le trafic ferroviaire en regardant les effets des perturbations sur le plan de transport.

Les simulateurs ferroviaires

L'enjeu de toute simulation est de reproduire la réalité de manière virtuelle et approchée à un coût moindre que l'expérimentation. En ferroviaire, l'expérimentation est justement très coûteuse. Le but de ces simulateurs est de reproduire le déplacement des trains sur une infrastructure selon le plan de transport que l'on veut tester. On distingue notamment les simulateurs selon leur échelle de description de l'infrastructure :

- Cette échelle peut être microscopique : on décrit alors l'infrastructure de manière très détaillée en prenant en compte tous les signaux et tous les aiguillages de manière exacte ainsi que les règles de comportement des trains à chaque instant.
- Elle peut également être macroscopique : on ne décrit plus les cantons de manière précise et on se contente de généralement de regarder ce qui se passe de gare à gare voire à des échelles encore plus grandes. Les règles auxquelles obéissent les trains à ces échelles ne sont plus les règles de comportement instantané bien qu'elles en découlent.

On distingue aussi les simulateurs sur un certain nombre d'autres caractéristiques dont voici une liste non-exhaustive. Aucun simulateur ne réunit toutes ces caractéristiques en même temps.

- La prise en compte ou non du calcul de la robustesse.
- La nature de la prise en compte du calcul des retards qui peut être stochastique ou déterministe.
- La nature des perturbations auxquelles on soumet le système.
- L'implémentation ou non d'une gestion de la réorganisation du trafic (dans le cas où le plan de transport n'est plus respecté)
- etc...

Les simulateurs ferroviaires demandent généralement l'ajustement précis d'un très grand nombre de paramètres pour fonctionner correctement. Les simulations se montrent parfois très sensibles à des variations petites de ces paramètres sans que l'on sache toujours l'interpréter. Cela limite généralement les conclusions que l'on peut en tirer. De plus les spécificités régionales et nationales des réseaux de chemin de fer ont conduit au développement d'un grand nombre de simulateurs qui obéissent à des spécificités nationales ou régionales et ne sont pas forcément facilement transposables d'un pays à l'autre. Aujourd'hui, la construction d'un plan de transport bien qu'elle puisse reposer sur l'usage de ces outils, repose encore beaucoup sur l'intelligence humaine. Dans la partie suivante nous allons justement décrire comment le système ferroviaire peut s'écartier de ce plan de transport.

1.2.5 Effondrement du plan de transport

Définition - Effondrement du plan de transport

On considère dans cette thèse qu'il y a un effondrement du plan de transport lorsqu'un retard apparaît sur une ligne, impactant plusieurs trains sans possibilité de ratrappage de ce retard. La résolution de cette situation demande l'intervention de la régulation du trafic.

Avec la croissance spectaculaire de la demande depuis une quarantaine d'années, le transport ferroviaire en région parisienne a atteint ses limites et un consensus se dégage

entre pouvoirs publics, usagers et opérateurs de transport pour considérer que l'état du réseau ne permet plus de répondre de façon satisfaisante aux besoins actuels. La saturation des lignes, la fréquence des incidents, les retards à répétition... font partie du quotidien de bon nombre d'usagers et donnent le sentiment d'une dégradation inexorable des conditions de transport. Une étude interne à Transilien-SNCF [BENTZ \[2014\]](#) réalisée sur des données collectées entre janvier et avril 2014 montre à ce propos qu'aux heures de pointe, environ la moitié des trains de la ligne B présentent des retards d'au moins cinq minutes.

L'importance des retards a conduit Transilien à comptabiliser 46 causes de retards possibles classées en 16 catégories. Parmi ces causes, on trouve entre autres : «erreur de conduite», «affluence voyageurs», «défaut technique du matériel roulant», «signal d'alarme», «mauvaises conditions météorologiques»... Cette catégorisation des retards est une bonne démarche mais leur inscription dans une base de données n'est pas systématique et comporte parfois des erreurs ce qui ne facilite pas l'analyse de l'effondrement du plan de transport.

La problématique d'un effondrement du plan de transport n'est pas spécifique à Transilien, on trouve d'ailleurs dans la littérature de nombreux articles qui illustrent la tendance des systèmes ferroviaires planifiés à s'écartez de leur plan de transport. Ce sujet est abordé sous différents angles dans cette littérature. Parmi les axes de recherches les plus fréquents, on peut distinguer les articles qui s'intéressent à la construction d'un plan de transport optimal (voir par exemple [HIGGINS et collab. \[1996\]](#); [LIEBCHEN et collab. \[2010\]](#)), ce qui constitue une gestion en amont du problème, de ceux qui étudient la possibilité d'une gestion optimum, par les régulateurs, d'une situation d'effondrement du plan de transport (voir par exemple [CACCHIANI et collab. \[2014\]](#); [GÉLY et collab. \[2006\]](#)) ce qui constitue une gestion en aval du problème.

Enfin, une voie de recherche qui nous intéresse particulièrement dans ce manuscrit c'est l'étude de la propagation des retards sur le réseau qu'elle soit empirique ou théorique. Une chronologie de la littérature autour de cet axe de recherche est présentée au début du chapitre 2. Il faut noter que ces différentes voies de recherches sont fortement liées les unes aux autres, en effet un algorithme de recherche du plan de transport optimal peut par exemple reposer sur un modèle théorique de propagation des retards.

Nous avons plusieurs fois comparé le routier et le ferroviaire depuis le début de ce chapitre. Malgré les différences entre les deux modes de transport ¹⁰, le trafic routier présente une similitude avec le trafic ferroviaire qui a retenu notre attention : l'existence d'un état du trafic dégradée que l'on appelle état congestionné en trafic routier. On va maintenant présenter quelques notions de trafic routier.

10. Ces différences seront caractérisées en détail dans la section suivante.

1.3 Le trafic routier et la congestion

La congestion est sans doute l'un des phénomènes les plus connus en trafic routier au point qu'il alimente de nombreuses conversations tant chez les usagers que chez les spécialistes. Cette omniprésence de la congestion dans le débat public pourrait faire croire qu'il s'agit d'une notion simple. Il n'en est rien : la congestion est un concept pour lequel il existe de nombreuses définitions qui ne sont pas toujours compatibles entre elles. Depuis environ soixante ans des chercheurs tentent de comprendre les mécanismes à l'œuvre dans ce phénomène. Nous allons présenter brièvement dans cette partie un modèle de trafic routier, le modèle LWR et le diagramme fondamental (DF) qui est un de ces composants. Nous allons en particulier nous intéresser au passage des variables individuelles aux variables collectives ainsi qu'aux hypothèses nécessaires à la construction d'un tel diagramme. Enfin nous illustrerons, à la lumière de ce DF, ce qu'est la congestion.

Le modèle LWR a été développé par [LIGHTHILL et WHITHAM \[1955\]](#); [RICHARDS \[1956\]](#). C'est un modèle qui a beaucoup d'avantages : il permet de reproduire le comportement du trafic routier avec peu de variables. Des situations parfois complexes peuvent être analysées de manière analytique ce qui rend le modèle plus facile à utiliser et les situations étudiées plus compréhensibles. De plus ce modèle fait le lien entre la réalité discontinue du trafic routier puisque constitué de briques élémentaires, les véhicules, et une vision continue du trafic, analogue à la mécanique des fluides, où le trafic est vu à travers des grandeurs continue comme le débit ou la concentration.

1.3.1 Variables individuelles-variables de flux

La vision naturelle du trafic n'est pas celle des milieux continus, c'est la vision d'un ensemble de véhicules roulant et évoluant sur des routes. Pour décrire cette vision les données adaptées sont les trajectoires des véhicules c'est-à-dire leur position au cours du temps $x(t)$. On peut alors obtenir les vitesses $v(t)$ en dérivant les positions $x(t)$ une fois et l'accélération $a(t)$ en les dérivant deux fois. Le plus souvent on doit se contenter expérimentalement d'une partie seulement de ces données, soit du fait que tous les véhicules ne sont pas associés à des données GPS récupérables, soit du fait que les capteurs ne mesurent le trafic qu'en certains points du réseau. Les informations manquantes peuvent alors être extrapolées si la qualité des informations disponibles le permet.

Une telle description microscopique du trafic est riche en information puisqu'elle consiste en une suite de positions pour les instants t . Cette richesse est à double tranchant car la manipulation d'un tel jeu de données n'est pas aisée de sa taille, ce qui rend l'analyse et la compréhension des phénomènes de trafic difficile.

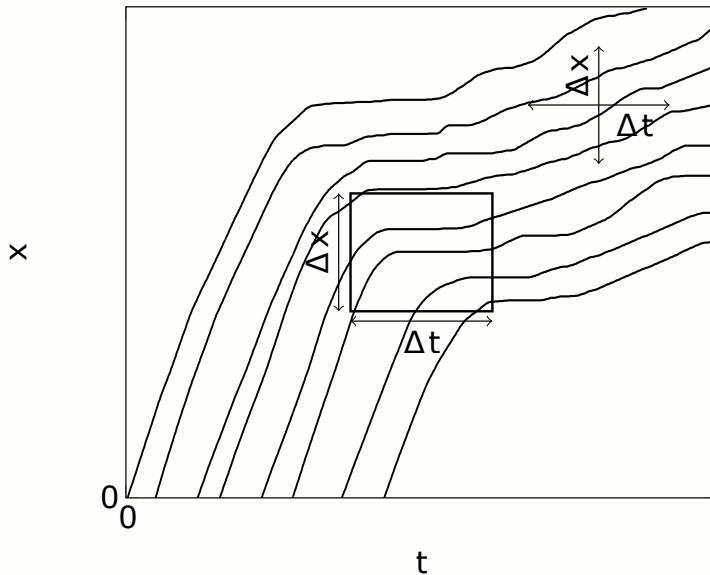
Au lieu de considérer les véhicules individuellement, on peut les regarder collectivement ; on procède de la façon suivante. On représente tout d'abord les trajectoires sur un graphique $x(t)$. Autour d'un point de ce graphique, on définit un pas de temps Δt et un pas spatial Δx . Il y a i véhicules qui traversent le segment Δx et j véhicules qui traversent le segment Δt aux vitesses v_j . On peut alors calculer un débit moyen et une concentration moyenne :

$$q(\Delta t) = \frac{i}{\Delta t} \quad (1.3)$$

$$k(\Delta x) = \frac{\sum_j v_j}{\Delta x} \quad (1.4)$$

Les segments Δx et Δt sont représentés par une croix sur la figure 1.9. Cette méthode de calcul présente deux problèmes : d'une part, les débits et les concentrations ne sont pas évaluées sur le même ensemble de données. D'autre part, les calculs du débit et de la concentration présentent des discontinuités fortes liées au faible nombre de véhicules qui composent ces moyennes.

Figure 1.9 - trajectoires de véhicules et zone d'Édie



Sur ce graphique on a représenté les trajectoires de quelques véhicules $x(t)$ ainsi que la zone d'Édie que l'on utilise pour le calcul du débit et de la concentration.

(source de l'image : CHIABAUT [2009])

EDIE [1963] et **DAGANZO** [1997] ont proposé des méthodes qui permettent de s'affranchir de ces deux problèmes pour évaluer la concentration et le débit. Le principe est de découper l'espace-temps en sous-région avec un pas de temps Δt et un pas spatial Δx . On appelle une telle sous-région une zone d'Édie. On peut alors associer à chaque zone d'Édie un débit moyen et une concentration moyenne qui sont définis de la manière suivante. Dans chaque sous-région, on note l_i la longueur et t_i la durée du trajet parcouru dans

cette région par le véhicule i . On note q le débit moyen et k la concentration moyenne¹¹.

$$q(\Delta t, \Delta x) = \frac{\sum_i l_i}{\Delta x \cdot \Delta t} \quad (1.5)$$

$$k(\Delta t, \Delta x) = \frac{\sum_i t_i}{\Delta x \cdot \Delta t} \quad (1.6)$$

La figure 1.9 représente quelques trajectoires de véhicules ainsi que la zone d'Édie qui sert aux calculs du débit et de la concentration. Quelle que soit la taille de ces sous-zones d'Édie, le nombre de véhicules restera toujours faible devant le nombre d'Avogadro qui sert de justification à l'hypothèse des milieux continus en mécanique des fluides. Il faut faire attention à ne pas prendre des zones d'Édie de l'espace-temps trop petites afin de garder une évolution continue des variables.

En pratique, la démonstration de la validité de l'hypothèse des milieux continus n'a jamais été faite en trafic routier et le reproche en est souvent fait (voir LECLERC [2009]). La justification de l'usage de cette hypothèse est empirique, on considère que si les variables q et k varient effectivement comme des variables continues alors, l'hypothèse des milieux continus est valable. Le trafic routier n'est pas le seul domaine des sciences pour lequel la démonstration du passage aux milieux continus est controversée : on utilise en physique statistique l'hypothèse d'ergodicité, or dans de nombreux cas cette hypothèse reste non démontrée (voir DIU et collab. [1989b]; SZÁSZ [1996]), et alors seule l'adéquation, souvent parfaite en physique statistique, entre les résultats expérimentaux et le calcul théorique semble valider cette hypothèse.

En réalité, la comparaison qui est souvent faite entre le nombre de voitures que l'on prend en compte dans une zone d'Édie ($25 \text{veh}.\text{km}^{-1}$) et le nombre d'Avogadro : $N_A = 6 \times 10^{23}$, cache deux erreurs de raisonnement :

- Ce n'est pas parce qu'un nombre de particules supérieur au nombre d'Avogadro peut servir à justifier l'hypothèse des milieux continus, qu'on doit penser qu'un nombre de particules même très inférieur au nombre d'Avogadro implique qu'on ne puisse pas utiliser l'hypothèse des milieux continus. C'est une erreur de logique. De plus d'après DIU et collab. [1989a], pour un volume de gaz de 1mm^3 à pression ambiante il n'y a «que» 10^{15} molécules de gaz et l'hypothèse des grands nombres reste parfaitement valable à cette échelle et même à des échelles encore inférieures. Certes ce nombre est encore grand devant le nombre de véhicules que l'on peut comptabiliser dans une zone d'Édie, mais un second argument revient mettre en cause le bien-fondé de cette comparaison récurrente.
- L'article DIU et collab. [1989a] nous fournit un bel exemple qui illustre que la comparaison est incomplète : quand le courant circule dans un fil électrique, la vitesse du mouvement d'ensemble des électrons de conduction est de l'ordre de grandeur du mm.s^{-1} . Mais l'ordre de grandeur de la vitesse de ces mêmes électrons pris individuellement est le km.s^{-1} . Et pourtant l'hypothèse des milieux continus s'applique

11. Ces notations sont usuelles en trafic routier, en ferroviaire on utilisera dans ce manuscrit d'autres notations et on le précisera

au courant électrique. En comparaison, en trafic routier, on trouve les mêmes ordres de grandeur pour la vitesse moyenne d'un ensemble de véhicules et la vitesse individuelle de ces véhicules.

En réalité toujours selon DIU et collab. [1989a] l'hypothèse des milieux continus se valide classiquement dans le cadre du théorème de la limite centrale¹². Appliquons ce théorème aux électrons dont on vient de parler. La vitesse de chaque électron suit une distribution centrée sur une valeur de l'ordre du $mm.s^{-1}$ qui a une largeur de l'ordre du $km.s^{-1}$ soit 10^6 fois plus. Le théorème de la limite centrale nous dit qu'il faudra considérer un ensemble de $(10^6)^2 = 10^{12}$ électrons pour que la largeur de la distribution de cet ensemble soit de l'ordre du $mm.s^{-1}$ c'est-à-dire le même ordre de grandeur que sa moyenne. Pour les modes routier et ferroviaire, les ordres de grandeur de la vitesse d'un véhicule d'une part, et de la largeur de la distribution de cette vitesse d'autre part, sont les mêmes. On a donc avec une seule particule une situation qui ressemblent à ce qu'on obtient avec 10^{12} électrons.

1.3.2 Relation entre vitesse, concentration et débit du flux.

Puisque l'on considère que l'hypothèse des milieux continus est valide, on reprend un résultat de la mécanique des fluides qui fait le lien entre la vitesse u , la concentration k et le débit q d'un flux :

$$q = k \times u \quad (1.7)$$

On peut donc estimer les vitesses moyennes d'une sous-région en calculant :

$$u = q/k$$

1.3.3 Diagramme fondamental

Si on recherche empiriquement à relier le débit et la concentration, on observe qu'il existe une concentration critique k_c qui sépare deux régimes de trafic différents. On peut le vérifier sur la figure 1.10.

- Pour les concentrations inférieures à la concentration critique, le débit croît avec la concentration. On dit que le trafic est fluide. On peut faire le lien avec une explication microscopique : pour ces concentrations, la distance inter-véhiculaire d'une paire de véhicules prise dans le flux est assez grande pour que le conducteur du véhicule suiveur considère qu'il peut rouler à la vitesse maximale autorisée u_{max} . Cela restant valable pour chacune des paires de véhicules, le coefficient de proportionnalité entre le débit et la concentration s'écrit :

$$q = k \times u_{max}$$

12. Attention aux hypothèses de ce théorème et en particulier à l'indépendance entre variables aléatoires.

- Pour les concentrations supérieures à la concentration critique, l'augmentation de la concentration entraîne une diminution du débit. À l'échelle microscopique, les distances inter-véhiculaires deviennent faibles et les conducteurs ralentissent alors pour préserver leur sécurité. On dit que le trafic est en congestion.

La représentation graphique de la relation que l'on décrit entre débit et concentration s'appelle le diagramme fondamental. Avec l'aide de la relation 1.7 on peut construire trois représentations de ce diagramme :

$$\begin{aligned} q &= F(k) \\ u &= F(q) \\ u &= F(k) \end{aligned}$$

On n'utilisera que la première représentation du diagramme dans ce manuscrit. On trace expérimentalement la relation $q = F(k)$ à partir des données sur une section autoroutière de la M6 Highway près de Manchester au Royaume-Uni et présentée dans CHIABAUT [2009]. On s'intéresse aux données issues d'un seul capteur le M7135. On ne dispose pas des trajectoires des véhicules et on ne peut donc pas évaluer le débit et la concentration avec les formules 1.3, 1.4, 1.5 et 1.6. Le capteur nous fournit toutes les minutes, le nombre de véhicules qui passent sur une section de voie pendant cette minute, ainsi que la vitesse moyenne de cet ensemble de véhicules sur cette section de voie. On a directement le débit en ($veh.min^{-1}$) et on peut évaluer la concentration grâce à la formule 1.7.

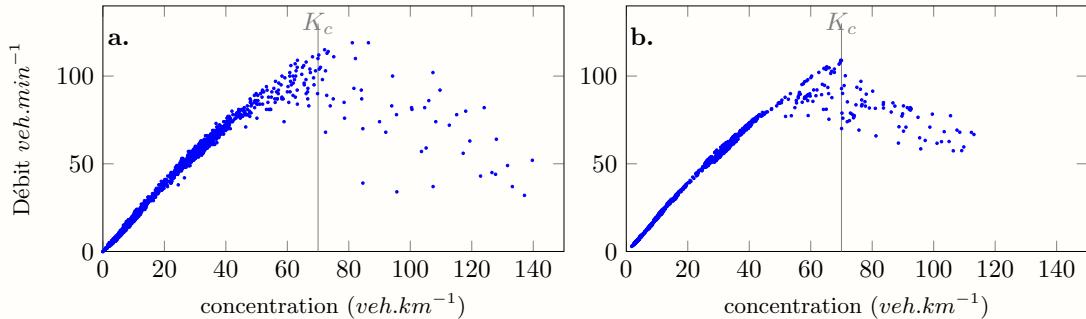
La figure 1.10 présente deux tracés de cette relation $q = F(k)$. Pour la sous-figure **a** on a une période d'agrégation des données d'une minute et pour la sous-figure **b** on a une période d'agrégation des données de dix minutes. On note effectivement assez nettement l'existence d'une concentration critique K_C qui sépare un état fluide du trafic à gauche, d'un état congestionné à droite. En revanche si cette distinction entre deux états est nette, la valeur frontière de k_c ne peut être déterminée précisément. On remarque aussi qu'augmenter la taille de la période d'agrégation permet de diminuer la dispersion, surtout dans la partie congestionnée du DF. Cette augmentation de la période d'agrégation a aussi un effet sur le débit maximum.

1.3.4 Le modèle LWR

Le modèle LWR fait l'hypothèse qu'il existe une relation concave entre le débit q et la concentration k que l'on écrit : $q = Q(k)$. Cette relation, qui décrit des états d'équilibre du trafic en une section d'un axe routier a été proposé en premier par GREENSHIELDS et collab. [1935], avec la proposition d'un DF parabolique. Mais la forme la plus simple mais aussi la plus répandue est le DF triangulaire proposé par BANKS [1989]; NEWELL [1993]. Ce DF est prisé car il décrit assez fidèlement les observations expérimentales et n'utilise que trois paramètres¹³ :

- u la vitesse des véhicules en situation fluide

13. Cette forme reste néanmoins controversée car elle échoue à expliquer certains phénomènes observés en trafic routier.

Figure 1.10 - Exemples de diagramme fondamental expérimental sur autoroute


Exemples de diagramme fondamental expérimental sur autoroute. La sous-figure **a** on a une période d'agrégation des données d'une minute et pour la sous-figure **b** on a une période d'agrégation des données de dix minutes. Cette plus grande période d'agrégation des données conduit à une dispersion moins grande des points sur le DF. On a représenté k_c la concentration critique, cependant la dispersion des points sur ces diagrammes ne permet pas de la placer précisément.

- k_{max} la concentration maximale
- w la vitesse des ondes de redémarrage.

Son équation s'écrit :

$$Q(k) = \min(u \cdot k, w \cdot (k - w)) \quad (1.8)$$

La vitesse des ondes de redémarrage, w , est négative car, lors du redémarrage d'une file de véhicules à l'arrêt, la propagation de l'onde de redémarrage se fait vers l'arrière. On représente un DF triangulaire sur la figure 1.11. On a dit que les points du diagramme fondamental, selon le modèle LWR, correspondaient à des zones d'équilibre stable du trafic. Nous allons maintenant préciser cette notion.

Tout d'abord, pour tracer la figure 1.10, nous n'avons pas pris de précautions particulières. Une partie de la dispersion des points est due au fait que nous avons affiché des points qui ne correspondent pas à des états d'équilibre du trafic.

Définition - Équilibre en trafic routier

Sur une zone d'Édie dans le graphe de la représentation des trajectoires, on est à l'état d'équilibre si les trajectoires sont rectilignes, parallèles et régulièrement espacées. (voir CHIABAUT [2009])

Par exemple la zone d'Édie sélectionnée sur la figure 1.9 ne correspond pas à un état d'équilibre du trafic. Cette définition pose un problème : en réalité les trajectoires des véhicules ne sont jamais parfaitement parallèles, il faut donc définir un critère quantitatif pour dire si une zone d'Édie correspond ou non à un état d'équilibre. CHIABAUT [2009] présente

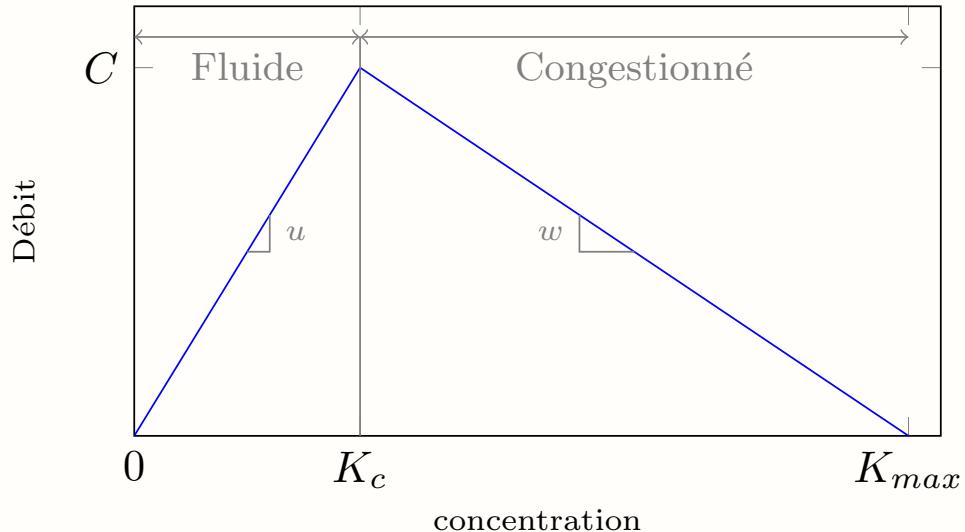
Figure 1.11 - Diagramme fondamental triangulaire


Diagramme fondamental triangulaire théorique. On distingue à gauche la partie fluide du DF avec une pente u qui correspond à la vitesse des véhicules en trafic fluide et à droite la partie congestionnée avec une pente $-w$ qui correspond à la vitesse des ondes de redémarrage du trafic. Le débit devient nul quand le flux atteint la concentration maximale k_{max} . Entre ces deux états du trafic, on trouve le point de débit maximal C pour une concentration k_c .

des méthodes d'analyse du trafic qui permettent de déterminer si l'on est à l'équilibre ou non. Cette hypothèse rejoint la discussion précédente sur l'approximation des milieux continus.

En effet, on a vu qu'à l'échelle microscopique la vitesse des électrons dans un fil conducteur était très grande devant la vitesse d'ensemble qui forme le courant. Pour faire la comparaison avec le trafic routier, on ne voit pas à cette échelle un ensemble de trajectoires $x(t)$ rectilignes et parallèles. Pour passer aux milieux continus, on divise l'espace en sous-volumes qui contiennent un grand nombre d'électrons mais ces espaces restent petits devant l'échelle du fil conducteur. Cette échelle intermédiaire s'appelle l'échelle mésoscopique¹⁴. Si on divise notre fil électrique en sous-volumes mésoscopiques et que l'on assimile ces sous-volumes à des particules, alors en traçant les trajectoires $\vec{X}(t)$ de ces pseudo-particules, on va effectivement observer qu'elles sont localement rectilignes et parallèles entre elles. Dans le modèle LWR, l'hypothèse d'équilibre du trafic est à relier avec le fait que les échelles microscopiques et mésoscopiques usuelles en physique statistique sont confondues en trafic routier.

14. Attention, le mot mésoscopique n'est pas toujours utilisée de la même manière en trafic routier et en physique statistique.

1.3.5 Congestion-définitions

On a annoncé au début de cette partie qu'il existait plusieurs définitions possibles pour la congestion en trafic routier, on va maintenant présenter les plus importantes d'entre elles.

Définition - Congestion-définition locale

Sur une section de voie, à partir du diagramme fondamental, on dit que le trafic est en congestion s'il évolue dans la partie droite du diagramme, c'est-à-dire si la concentration est supérieure à la concentration critique.

Définition - Congestion-économie des transports

Un système est congestionné si l'ajout d'un véhicule supplémentaire dégrade le niveau de service pour les autres.

Ces deux premières définitions sont compatibles entre elles, en effet lorsque l'on franchit la concentration critique sur le diagramme fondamental en ajoutant un véhicule dans le flux, on diminue la vitesse de tous les véhicules et on rallonge leur temps de parcours. On peut également définir la congestion à partir de l'offre et de la demande.

Définition - Offre et demande en trafic routier

L'offre correspond aux possibilités d'écoulement de l'infrastructure. La demande est le trafic qui cherche à s'écouler et qui s'écoulerait s'il n'était pas contraint.

Ces deux notions ont du sens tant à l'échelle locale en un point précis de l'infrastructure, qu'à l'échelle d'une portion de réseau toute entière.

Définition - Congestion-offre et demande

En situation de congestion, la demande n'est pas satisfaite, c'est l'offre qui conditionne le débit. En situation de trafic fluide, au contraire, toute la demande est satisfaite, c'est elle qui fixe le débit.

1.3.6 Comparaison entre le trafic ferroviaire et le trafic routier

On vient de présenter brièvement dans les paragraphes précédents le cadre théorique autour duquel s'est développé la science du trafic routier au xx^e siècle. Un certain nombre

des concepts présentés interpelle quant à leur possible transposition au domaine du ferroviaire. Nous allons livrer ici un comparatif méthodique en vue de déterminer ce que la théorie du trafic routier pourrait apporter au monde du ferroviaire et les obstacles à surmonter pour parvenir à de telles fins.

L'hypothèse des milieux continus

On a vu que l'hypothèse des milieux continus restait une question ouverte en théorie du trafic routier. Néanmoins on peut considérer que cette hypothèse est empiriquement justifiée dans un grand nombre de situations en trafic routier. L'ordre de grandeur de la concentration routière en véhicules est de $25\text{veh}.\text{km}^{-1}.\text{voie}^{-1}$, on a vu que ces concentrations, plutôt faibles en comparaison avec ce qu'on trouve habituellement en physique statistique, ne sont pas forcément une limitation pour développer une vision continue du trafic routier. Mais comme toute vision continue, elle perd inévitablement sa signification en dessous de l'échelle microscopique, c'est-à-dire en dessous de $40m$ pour une concentration de $25\text{veh}.\text{km}^{-1}.\text{voie}^{-1}$. De même en ferroviaire l'ordre de grandeur de la concentration est de $1\text{veh}.\text{km}^{-1}.\text{voie}^{-1}$, si vision continue il y a, elle ne pourra pas prétendre décrire la réalité en dessous de l'échelle du kilomètre.

Les états d'équilibres

Le tracé des trajectoires des véhicules $x(t)$, sur un diagramme spatio-temporel constitue un outil qui est utilisé tant en routier qu'en ferroviaire. On a vu qu'en routier l'identification des états d'équilibre pouvait se faire sur ce diagramme par identification (visuelle en première approche) des zones pour lesquelles les trajectoires des véhicules sont à peu près parallèles, rectilignes et régulièrement espacées. Pour qu'une zone d'Édie puisse remplir une telle condition il faut donc qu'elle contienne au moins trois trajectoires. Du fait de la faible densité du trafic ferroviaire, la recherche de zones d'équilibre ne doit se faire que sur des zones d'Édie dont le pas spatial est au minimum d'environ $3km$.

L'interaction entre véhicules

Les bancs de poissons, les vols d'étourneaux, les foules, le trafic routier et le trafic ferroviaire ont plusieurs choses en commun. Tout d'abord les particules qui les composent sont auto-propulsées, elles n'ont pas besoin d'être soumises à l'influence des autres particules pour pouvoir se mouvoir. La seconde caractéristique commune réside dans l'interaction particulière qui régit les mouvements collectifs de ces particules : cette interaction est basée sur la vue qu'une de ces particules a d'une autre. Or, la vue est anisotrope, elle donne une orientation à chaque particule, un devant et un derrière. Cela entraîne que l'interaction entre les particules est asymétrique : pour reprendre l'exemple du trafic routier, entre un véhicule suiveur et un véhicule suivi, l'interaction implique que le premier est influencé

par le second sans que la réciprocité ne soit vraie. On peut modéliser ces interactions par des forces mais ces pseudos-forces violent alors la troisième loi de Newton¹⁵.

Le trafic ferroviaire est un cas particulier quant à la nature de cette interaction : comme on l'a vu en 1.1.2 les distances de freinage nécessaires pour assurer la sécurité des trains sont supérieures à la distance de visibilité. Il a donc fallu prolonger cette vision, ce qui est le rôle du cantonnement. Cette prolongation s'est faite de manière discrétisée tant spatialement par l'implantation des signaux sur la voie qu'au niveau de la valeur de l'interaction qui ne peut prendre que trois valeurs. Un état de l'art des récents développements dans le domaine des particules auto-propulsées peut être lu dans **VICSEK et ZAFEIRIS [2012]** ou dans **HELBING [2001]** qui est plus orienté vers les mouvements de foule et le trafic routier.

Les degrés de liberté

On a vu en 1.1.2 que le guidage des trains impliquait pour eux un degré de liberté moindre en comparaison avec les voitures. Cela implique qu'il est généralement plus difficile pour un train de dépasser que pour une voiture. Mais en réalité il faut aussi prendre en compte la planification en ferroviaire car elle constraint encore plus le mouvement des trains qui n'interagissent pas seulement entre eux mais aussi avec un horaire.

Il existe également des situations pour lesquelles le train interagit avec un autre acteur du système ferroviaire. Le trafic est contrôlé en temps réel par des opérateurs de régulation du trafic qui peuvent volontairement modifier le plan de transport pour faire face à une situation dégradée. Le système ferroviaire est donc piloté à deux niveaux : les conducteurs pilotent les trains avec une marge de manœuvre faible tant que le plan de transport est respecté pour l'ensemble des trains. La régulation du trafic peut si le système s'écarte du plan de transport à l'instant t , modifier le plan de transport à l'instant $t + 1$ pour espérer rétablir une situation, la plus proche possible du plan de transport à l'instant $t + 2$.

L'offre et la demande

Quel que soit le mode de transport, routier ou ferroviaire, il est soumis aux lois économiques de l'offre et de la demande. Mais ces deux modes de transport divergent dans leur approche de la gestion de l'offre et de la demande. En ferroviaire c'est le plan de transport qui est construit en fonction d'une étude de la demande en transport de la population. Une fois ce plan de transport construit, on change de vision : sous l'angle de la comparaison entre l'offre et la demande on va analyser le déroulement d'un épisode de trafic (mod) en ferroviaire on va comparer ce qui était planifié avec ce qui a été effectivement fait. On a résumé dans un tableau : 1.1, quelques différences entre le trafic ferroviaire urbain et la circulation automobile urbaine sur une autoroute urbaine ou sur un réseau de rues.

15. En réalité, la troisième loi de Newton qui dit que $\vec{F}_{A/B} = -\vec{F}_{B/A}$ est bien sûr respectée. La véritable force physique à lieu entre la particule et le support sur lequel elle s'appuie, le rail par exemple.

Tableau 1.1 - comparatif ferroviaire-routier

Variables	chemin de fer urbain	autoroute urbaine	réseau routier
vitesse	$< 140 km.h^{-1}$	$< 130 km.h^{-1}$	$< 50 km.h^{-1}$
longueur des trajets	100km	20km	20km
densité	1 veh.km^{-1}	$25 \text{ veh.km}^{-1}.lane^{-1}$	$25 \text{ veh.km}^{-1}.lane^{-1}$
accélération ($m.s^{-2}$)	$-0.7 < a < 0.5$	$-3.5 < a < 2.5$	$-3.5 < a < 2.5$
interaction entre véhicules	par la signalisation	par contact visuel	par contact visuel et feux de signalisation
dépassagement	rare	fréquent	fréquent
arrêts	en gare ou sur signal	pas d'arrêt	sur signal
itinéraire	= 1 par O/D* fixé à l'avance	1 to 3 par O/D*	$\gg 1$ par O/D*
maillage	nul ou faible	moyen	haut
capteurs	temps de passage	temps de passage et vitesse	temps de passage et vitesse
précision des capteurs	± 1 minute	$\pm 15\%(\text{veh}.min^{-1})$	$\pm 15\%(\text{veh}.min^{-1})$
planification	globale	individuelle	individuelle

*O/D veut dire matrice origine destination

1.4 Problématique

Cette thèse s'inscrit donc dans un contexte caractérisé par des insuffisances du système de transport ferroviaire en région parisienne qui ont amené et amènent encore Transilien à remettre en cause ses méthodes de travail dans le but de réaliser un saut de performance. Cette ambition ne peut pas reposer à court terme sur l'amélioration d'une infrastructure vieillissante, amélioration d'autant plus lente qu'elle coûte cher et qu'elle est ralentie par de mauvaises perspectives économiques. Dès lors, une meilleure gestion des ressources existantes devient nécessaire afin d'atteindre l'objectif fixé.

Une meilleure compréhension des phénomènes par lesquels le plan de transport se dégrade, provoquant au passage l'insatisfaction des clients, est nécessaire comme base pour développer des outils ou des méthodes pour agir aussi bien en amont des problèmes, lors de la construction du plan de transport, qu'en aval, lors de la régulation du trafic.

L'utilisation de simulateurs de trafic ferroviaire au sein de Transilien-SNCF a son utilité, mais leur utilisation n'est pas facile. Ces outils, comme vu précédemment, demandent l'ajustement précis d'un grand nombre de paramètres pour pouvoir fonctionner et se montrent très sensibles à des variations de ces mêmes paramètres. S'ils sont utilisés correctement, ils peuvent être adaptés pour étudier une situation précise d'un réseau ferroviaire mais leur contribution à la compréhension globale du système est faible.

Il y a donc un besoin d'utiliser des outils de visualisation et de modélisation du trafic plus simples à manipuler et donc plus simples à comprendre mais avec une ambition intacte d'observer et de reproduire les phénomènes impliqués dans la dégradation du plan de transport. De tels outils existent déjà en trafic routier et malgré les différences qui existent entre les deux modes de transport, on se demande dans cette thèse si l'existence d'un état dégradé du trafic, en routier comme en ferroviaire, ne constitue pas une piste à suivre pour importer de tels outils pour l'étude du trafic ferroviaire.

Par ailleurs, la diversité des éléments de perturbation d'un plan de transport dans la réalité, ainsi que la non-prédictibilité d'un certain nombre d'entre eux nous a conduit à considérer l'hypothèse du caractère aléatoire ou stochastique de ces éléments perturbateurs. On se demande donc également dans cette thèse si la prise en compte de perturbations stochastique dans un modèle simplifié du trafic ferroviaire peut reproduire les phénomènes de dégradation du trafic observés dans la réalité.

Les travaux présentés dans ce manuscrit s'intéressent uniquement au cas d'un trafic ferroviaire planifié en zone urbaine dense et avec un cantonnement fixe à trois aspects.

Chapitre 2

Modélisation microscopique stochastique sans gare

Sommaire

2.1 État de l'art de la modélisation stochastique du déplacement des trains	44
2.1.1 Historique de la modélisation stochastique du déplacement des trains	44
2.2 Construction du modèle	49
2.2.1 L'infrastructure	49
2.2.2 Le matériel roulant	49
2.2.3 Les forces décrivant le comportement d'un train	49
2.2.4 Implémentation numérique du modèle	54
2.3 Apparition de la congestion	55
2.3.1 Organisation des simulations	55
2.3.2 Compétition entre la force aléatoire et la force horaire	56
2.3.3 Résultats	58
2.4 Analyse du modèle et des résultats par des processus de Markov	62
2.4.1 Processus de Markov	62
2.4.2 Les distributions de type phase	63
2.4.3 Lien avec le modèle	68
2.5 Quantification de la congestion	73
2.5.1 Organisation des simulations	73
2.5.2 Dynamique de la vitesse instantanée moyenné sur les trains	73
2.5.3 Diagramme Fondamental de Ligne Ferroviaire	75

2.1 État de l'art de la modélisation stochastique du déplacement des trains

Dans la littérature académique qui traite du sujet des retards dans les systèmes ferroviaires, on distingue généralement les retards primaires et les retards secondaires. Le caractère primaire du retard signifie qu'il est intrinsèque au train et ne résulte pas d'une interaction avec un autre train. Les retards primaires peuvent résulter de nombreuses causes qui sont évoquées dans le chapitre 1 mais n'ont pas pour cause un autre retard. C'est ce qui les distingue des retards secondaires qui sont causés par la propagation d'un retard primaire ou secondaire à travers le réseau. Une approche de modélisation déterministe implique ici de connaître des lois de prédiction déterministes pour les différentes causes de retards primaires ce qui paraît impossible : comment savoir à l'avance qu'un conducteur va être distrait et rouler plus lentement que prévu pendant une minute ? Comment savoir à l'avance le moment précis où un passager va retenir les portes du train et retarder son départ ? Dans ces conditions, une étude déterministe des retards paraît bien illusoire. Henri Poincaré écrivait déjà au début du siècle précédent dans **POINCARÉ [1908]**.

“Vous me demandez de vous prédire les phénomènes qui vont se produire. Si, par malheur, je connaissais les lois de ces phénomènes, je ne pourrais y arriver que par des calculs inextricables et je devrais renoncer à vous répondre ; mais, comme j'ai la chance de les ignorer, je vais vous répondre tout de suite. Et, ce qu'il y a de plus extraordinaire, c'est que ma réponse sera juste.”

Dans cette citation, Henri Poincaré défend l'usage du calcul probabiliste en physique. Face à l'ignorance d'un grand nombre de paramètres pourtant nécessaires à une modélisation déterministe des phénomènes physiques, il prône l'abandon de l'approche déterministe pour une approche basée sur le calcul des probabilités qui nécessite d'avoir moins d'informations sur le système étudié.

Le terme de retard n'est pas assez général pour l'usage que l'on va en faire dans ce manuscrit car il exclut les situations où il y a une avance. Nous allons préférer le terme d'écart à l'horaire dans la suite du manuscrit quand il s'agira de notre travail.

2.1.1 Historique de la modélisation stochastique du déplacement des trains

Depuis la première publication en 1974 qui traite de la nature stochastique des retards dans le monde ferroviaire jusqu'à aujourd'hui, on constate une continuité intellectuelle : des observations, répétées sur différents réseaux ferroviaires, ont mené à établir la forme de la distribution temporelle des retards à l'entrée d'une section de voie. On trouve dans la bibliographie différentes propositions d'approches autant numériques qu'analytiques qui utilisent ce résultat comme brique de base pour calculer de manière probabiliste la propagation des retards sur un réseau.

Le tableau 2.1 présente la chronologie de ce continuum intellectuel. On fait une distinction entre les travaux qui traitent de la distribution des retards et ceux qui traitent de

leur propagation¹. SCHWANHÄUSSER [1974]; WEIGAND [1981] furent les premiers à établir (publication en allemand) que la forme de la distribution temporelle des retards négatifs² obéissait à une loi exponentielle. ENGELHARDT-FUNKE et KOLONKO [2002] en ont fourni une preuve théorique en 2002. Ces résultats ont également été confortés par d'autres preuves expérimentales comme avec YUAN [2006].

Comme on vient de le dire, ce résultat a ensuite été utilisé comme brique de base pour modéliser la propagation des retards. MÜLHANS et WARNINGHOFF [1990]; WEIGAND

Tableau 2.1 - Chronologie du stochastique en ferroviaire

distribution des retards		propagation des retards	
1974	• Schwanhausser (A)		
1981	• Weigand (A)	1981	• Weigand (A)
		1990	• Mulhans (A)
		1994	• Carey & Kwiecinski
		1995	• Carey & Kwiecinski
1998	• Higgins & Kozan		
2002	• Engelhardt-Funke & Kolonko	2002	• Yuan & Goverde
2006	• Yuan (manuscrit de thèse)	2006	• Yuan (manuscrit de thèse)
		2007	• Meester & Muns
		2011	• Berger & Gebhardt
		2012	• Bucker & Seybold

Chronologie des papiers traitant de la modélisation stochastique des retards en ferroviaire. Les papiers de la première colonne étudient la distribution temporelle des retards des trains à l'entrée d'une section de voie donnée. Tous les papiers de cette colonne font état de recherche empirique à l'exception de ENGELHARDT-FUNKE et KOLONKO [2002] qui est théorique. Les papiers de la seconde colonne étudient la propagation de ces retards à partir de différentes hypothèses sur la façon de représenter mathématiquement la causalité d'un retard sur un autre et sur la façon de calculer cette propagation des retards. (A)signifie que la publication est écrite en allemand.

1. Il est tout à fait possible que les deux aspects soient abordés dans la même publication.

2. un écart à l'horaire (ou retard) négatif correspondant à une situation de retard selon la convention choisie dans cette thèse.

[1981] et HIGGINS et KOZAN [1998] ont proposé un modèle pour l'apparition de ces retards mais en faisant une hypothèse trop simplificatrice dans leurs calculs : le lien de causalité pouvant exister entre deux retards a été négligé. En d'autres termes on a considéré tous les retards comme étant indépendants les uns des autres il ne s'agit donc pas à proprement parler d'un modèle de propagation des retards. L'article de CAREY et KWIECIŃSKI [1994] s'attaque à cette lacune en considérant la succession de deux trains sur une même voie pour réintroduire le lien de cause à effet et donc la notion de propagation du retard. Dans CAREY et KWIECIŃSKI [1995] le raisonnement continue avec la découverte des relations de récursivité permettant d'exprimer le retard de n'importe quel train présent sur le réseau. Cette relation s'écrit :

$$\begin{bmatrix} \text{temps } t_i \text{ d'un} \\ \text{événements } i \end{bmatrix} = R_i \left(\begin{bmatrix} \text{événements} \\ \text{passés} \end{bmatrix}, \begin{bmatrix} \text{grille} \\ \text{horaire} \end{bmatrix}, \begin{bmatrix} \text{variables} \\ \text{aléatoires} \end{bmatrix} \right) \quad (2.1)$$

En étudiant les propriétés mathématiques de la fonction R_i , les auteurs établissent qu'elle doit représenter un processus de Markov continu en temps et à états continus. En pratique les équations de ce modèle sont difficiles à résoudre à cause de la description continue des états. Les publications MEESTER et MUNS [2007] puis BÜKER et SEYBOLD [2012] présentent un moyen d'approximer ce modèle avec une description discrète des états du modèle dans le processus de Markov : les processus de Markov à distribution de type phase (DTP).

Les processus de Markov DTP introduisent une alternative analytique à la traditionnelle approche de la méthode de Monte-Carlo pour résoudre un problème défini avec un processus de Markov. L'avantage des méthodes utilisant les DTP est de permettre d'évaluer les distributions des retards avec seulement une seule résolution numérique, alors qu'un grand nombre de résolutions numériques sont nécessaires dans la méthode de Monte-Carlo ce qui en fait une méthode particulièrement lente. En revanche, la méthode de Monte-Carlo tend à rassurer les tenants d'une vision ferroviaire plus classique qui pourront à chaque instant et pour chaque résolution numérique connaître l'état exact du système ferroviaire et identifier cet état à un cas réel, ce qui n'est pas possible avec des probabilités de retard. Cependant il convient de nuancer l'avantage apporté par les méthodes utilisant les DTP. En effet si les chaînes de Markov sont très longues et en l'absence de simplification calculatoire il est parfois impossible de résoudre un problème posé en termes de DTP. À l'inverse, la méthode de Monte-Carlo bien que lente est toujours opérationnelle et souvent plus simple à mettre en œuvre. En pratique, les méthodes de Monte-Carlo sont donc toujours utilisées aujourd'hui pour étudier la propagation des retards comme c'est le cas dans YUAN et collab. [2002].

Le but de tous les papiers présentés ci-dessus est d'améliorer la fidélité des outils de simulation existants. L'objectif de tels outils est la reproduction de la réalité avec un maximum de réalisme. Ce qui a un coût : les caractéristiques de l'infrastructure comme celles du matériel roulant doivent être décrites de la manière la plus précise possible. Alors même que l'essence de la citation de Henri Poincaré semble respectée grâce à l'introduction d'un ingrédient stochastique dans les modèles développés, la volonté, affichée dans ces publications, est de coller au plus proche d'une situation réelle. Cela demande la prise en

compte d'un grand nombre de paramètres qui sont spécifiques à un cas d'étude et ne permettent pas de généraliser les résultats à d'autres situations . Il en découle que les modèles développés servent uniquement à comprendre le fonctionnement d'un réseau ferroviaire en particulier et non à déterminer des principes généraux sur l'évolution stochastique des retards dans les réseaux ferroviaires en général.

De manière plus précise, le but des papiers présentés n'est pas tant de comprendre le phénomène de congestion ferroviaire, avec son apparition et sa propagation mais plutôt de faire en sorte que le modèle développé puisse reproduire statistiquement les retards observés sur un système réel afin d'être implémenté dans un simulateur pour servir d'outil opérationnel de construction de la grille horaire ou d'aide à la gestion de la circulation en temps réel. Le problème étant que les simulateurs ferroviaires sont loin d'être universels (voir chapitre 1) et sont en pratique souvent développés pour correspondre à des spécificités de réseaux nationaux ou régionaux. L'idée de la possibilité d'un modèle plus universel en ferroviaire n'est pas une idée répandue. [LAI et collab. \[2012\]](#) propose justement une nouvelle méthode de calcul de la capacité qui vise à cette universalité. Ce papier défend la possibilité que les travaux présentés puissent un jour aboutir à un modèle de trafic ferroviaire universel comme ce fut le cas pour le trafic routier quand le modèle LWR fut découvert et présenté dans l'article [LIGHTHILL et WHITHAM \[1955\]; RICHARDS \[1956\]](#).

La méthode reste cependant assez différente de l'approche abordée dans cette thèse : [LAI et collab. \[2012\]](#) constatent que la capacité ferroviaire donnée en nombre de trains par unité de temps ne tient pas compte de l'hétérogénéité du matériel. Il propose de rendre le calcul de la capacité indépendant du type de circulation en attribuant à chaque type de circulation un nombre de trains équivalent. À l'inverse, dans cette thèse, nous considérons des situations avec des trafics assez homogènes.

Le modèle présenté dans la partie suivante s'intègre dans une démarche qui repose sur la même ambition. Avec la simplicité de notre modèle et l'absence d'un travail calibration, on ne vise pas l'intégration de nos résultats dans un outil de simulation opérationnel. Au contraire, nous proposons de regarder la propagation des retards et surtout la résultante néfaste qu'est la congestion dans un cadre très général. Suite à des discussions avec des experts opérationnels de la gestion du trafic chez Transilien, nous avons décidé de nous intéresser en particulier à l'influence de trois paramètres sur l'apparition et la quantification de cette congestion : la vitesse maximum autorisée, c'est-à-dire la marge de vitesse autorisée au-dessus de la vitesse prévue par un plan de transport, l'amplitude des variations de la force aléatoire, et le nombre de trains.

Notations

Le tableau 2.2 présente les notations des principales variables utilisées dans ce chapitre. Il précise également quand cela est nécessaire les valeurs ou la plage de valeur que peuvent prendre ces variables par la suite. Ce tableau n'est pas exhaustif et d'autres variables secondaires peuvent être introduites lorsque cela sera nécessaire pour clarifier un calcul ou un raisonnement.

Tableau 2.2 - Notations

Nom	Valeur	Description
C	40	nombre de cantons
T	10-14	nombre de trains
F_{inter}	...	force d'interaction entre 2 trains
$F_{v \max}$...	force de sécurité
$F_{horaire}$...	force horaire
$F_{elastiq}$...	composante élastique de $F_{horaire}$
$F_{friction}$...	composante de frottements de $F_{horaire}$
F_{rand}	...	force aléatoire
i	...	index des trains
$plan$...	index accompagnant une variable planifiée
Δh	...	écart horaire d'un train
S_i	...	aspect du dernier signal croisé
$d(i, j)$...	distance entre l'avant de 2 trains
$d_{i,RS}$...	distance entre un train et le prochain signal rouge
P	3000kW	puissance de traction
m	200T	masse d'un train
l_t	200m	longueur d'un train
l_c	1000m	longueur d'un canton
l_{secu}	100m	distance de freinage face à un signal rouge
v_i	0-110 km.h ⁻¹	vitesse du train i
v_{max}	86-110 km.h ⁻¹	vitesse maximale autorisée
v_{safe}	30km.h ⁻¹	vitesse de sécurité
v_{plan}	85km.h ⁻¹	vitesse planifiée
a_{max}	0.6m.s ⁻²	maximum de l'accélération ou du freinage
$a_{elastiq}$	0.52m.s ⁻²	coefficient d'accélération de $F_{elastiq}$
$a_{friction}$	0.076m.s ⁻²	coefficient d'accélération de $F_{elastiq}$
$a_{r \max}$	$\pm(0.05 \text{ à } 0.3) \cdot a_{max}$	extrema de F_{rand}/m
t_{car}	90s	temps de réponse caractéristique de $F_{horaire}$
h_d	180s	temps que prends $F_{elastiq}$ pour saturer
$t_{friction}/h_d$	6,97	pente de tanh dans l'expression de $F_{friction}$
t_{corr}	300s	temps d'auto-correlation de F_{rand}

2.2 Construction du modèle

Dans cette partie, nous présentons une modélisation d'un système ferroviaire basée sur une infrastructure simplifiée, un matériel roulant homogène et une description du comportement de chaque train reposant sur quatre forces dont l'une d'entre elles est stochastique et permet de rendre compte des phénomènes aléatoires.

2.2.1 L'infrastructure

Considérons une ligne de train orientée dans un seul sens et bouclée sur elle-même. Cette ligne est constituée de C cantons associés à une signalisation à trois aspects telle qu'introduite dans la partie 1.1.2 à une petite différence près : on fait l'hypothèse simplificatrice que le temps de visibilité d'un signal est nul. Pour simplifier l'infrastructure, on fait également les hypothèses suivantes :

- Le terrain est plat.
- Tous les cantons ont la même longueur l_c .
- Tous les cantons admettent la même vitesse limite v_{max} qu'un train n'a pas le droit de dépasser. La longueur des cantons limite cette vitesse : un train qui rentre dans un canton à la vitesse v_{max} doit avoir le temps de freiner jusqu'à la vitesse de sécurité v_{secu} avant la fin de ce canton.
- L'infrastructure est fait de tel sorte qu'il est possible de définir un plan de transport qui, s'il est strictement respecté, prévoit que les trains traversent tous les cantons à la même vitesse v_{plan} . Cette vitesse est évidemment inférieure à la vitesse limite.

2.2.2 Le matériel roulant

Considérons T trains circulant sur la ligne que l'on vient de décrire. On fait l'hypothèse que tous les trains ont les mêmes caractéristiques physiques.

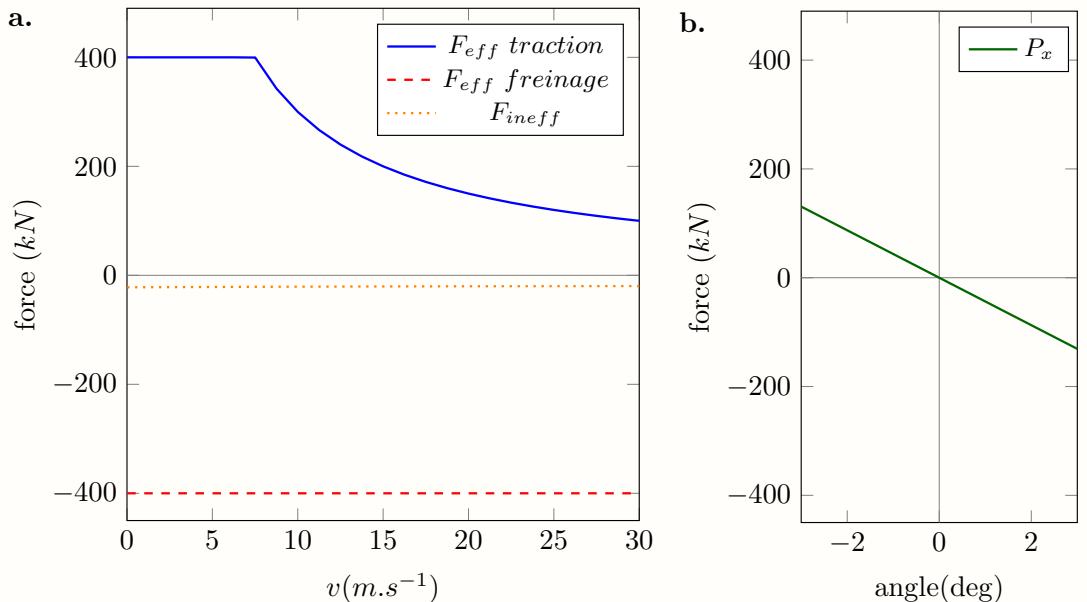
- Tous les trains font la même longueur l_t .
- Les accélérations (ou le freinage) des trains sont limités en valeur absolue par une accélération limite a_{max} . L'accélération peut prendre toutes les valeurs entre ces deux limites (voir discussion sur ce point en 2.2.3).
- Pour s'arrêter avant un signal rouge, un train roulant à v_{secu} commence à freiner à la distance de sécurité l_{secu} du signal rouge.
- Le comportement des trains est déterminé par un ensemble de quatre forces que nous allons présenter dans la partie suivante.

2.2.3 Les forces décrivant le comportement d'un train

On a présenté dans la partie 1.1.1 le bilan des forces qui s'appliquaient sur chacun des trains dans la figure 1.1 et dans le formulaire 1.1. Une application numérique nous a permis de représenter le graphe de ces forces dans la figure 2.1. On rappelle que F_{eff} représente les forces dites efficaces qui jouent un rôle utile et F_{ineff} représente les forces

dites inefficaces qui jouent un rôle pénalisant. C'est cette dernière figure qui va nous être utile pour justifier la pertinence d'une hypothèse que nous allons faire sur les valeurs que peuvent prendre l'accélération et le freinage des trains. Il n'est en effet pas rare de modéliser le déplacement d'un train en ne tenant compte que d'une capacité à accélérer et d'une autre à freiner sous la forme de deux valeurs d'accélération, l'une positive, l'autre négative, toutes deux indépendantes de la vitesse du train.

Figure 2.1 - Graphe des composantes de la force appliquée au train



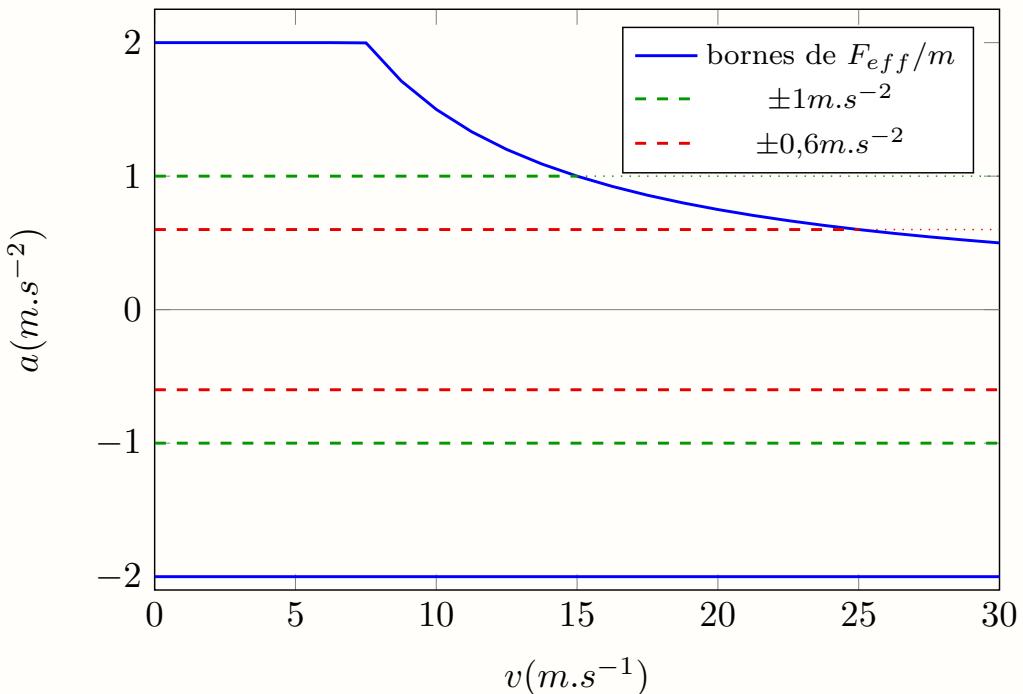
Cette figure, basée sur le formulaire 1.1 représente les graphes de l'expression des forces F_{eff} , F_{ineff} et P_x en fonction de la vitesse et de l'inclinaison. a. F_{eff} peut évoluer entre ses deux limites maximum de traction et minimum de freinage. F_{ineff} est un exemple de ce que pourrait valoir cette force au maximum (en valeur absolue). b. La force de déclivité ne dépend pas de la vitesse et est représentée en fonction de l'inclinaison.

Tentons justement de justifier la pertinence de cette hypothèse dans le cadre de la figure 1.2. On se situe dans le cadre des trains opérant en milieu urbain et dépassant rarement la vitesse de $30m.s^{-1} \approx 110km.h^{-1}$. Dans le cadre de trains roulant à des vitesses supérieures, les forces de frottements aérodynamiques deviendraient importantes et changeraient la forme de la courbe F_{ineff} présentée. Une valeur typique de l'accélération comme de la décélération maximale d'un train opérant en milieu urbain sur un terrain plat est de $\pm 1m.s^{-2}$. Mais ces valeurs sont rarement atteintes car elles participent à l'usure prématuée du matériel et à l'inconfort des passagers ; une bonne approximation est de considérer que l'on limite cette accélération à $\pm 0,6m.s^{-2}$.

On considère que F_{ineff} est suffisamment faible pour que les capacités d'un train à freiner et à accélérer sur du plat soient bien représentées par F_{eff} . Sur la figure 2.2 on a tracé F_{eff}/m qui représente l'accélération maximale autorisé a priori par la force F_{eff} , ainsi que les accélérations de $\pm 1m.s^{-2}$ et de $\pm 0,6m.s^{-2}$. On remarque que la force F_{eff}

semble pouvoir autoriser le train à avoir une accélération de $\pm 2 m.s^{-2}$. En réalité il n'en n'est rien : comme indiqué dans la partie 1.1.1, les mesures qui préviennent le patinage et l'enrayage nécessitent de ne pas s'approcher trop des limites d'adhérences car elles ne sont pas précisément établies. On remarque qu'en accélération l'approximation $a \leq 1 m.s^{-2}$ reste valable pour des vitesses jusqu'à environ $15 m.s^{-1} \simeq 60 km.h^{-1}$. De même l'approximation $a \leq 0,6 m.s^{-2}$ reste valable pour des vitesses jusqu'à environ $25 m.s^{-1} \simeq 90 km.h^{-1}$. Pour les besoins de la modélisation que l'on construit, on va considérer que l'ensemble des forces déterministes que l'on applique au train ne doivent pas conduire à des accélérations en dehors des bornes $\pm 0,6 m.s^{-2}$. L'ajout d'une force aléatoire va permettre de dépasser ces valeurs de 25%. Dans la suite de ce manuscrit, il ne sera pas rare de voir des forces présentées sous la forme d'un produit entre une masse et une accélération.

Figure 2.2 - Bornes de l'accélération d'un train



Cette figure permet d'évaluer le domaine de validité des deux hypothèses $|a| < 1 m.s^{-2}$ et $|a| < 0,6 m.s^{-2}$ par rapport aux limites autorisées par la F_{eff}/m . La première approximation cesse d'être valide en traction pour des vitesses supérieures à $15 m.s^{-1} \simeq 60 km.h^{-1}$ alors que la seconde l'est encore jusqu'à des valeurs de $25 m.s^{-1} \simeq 90 km.h^{-1}$.

La description du comportement du train sous forme de forces qui va suivre est basée sur l'observation d'un train réel dans différentes situations pour lesquelles on se pose la question suivante : quelle serait l'expression de la force qu'il faudrait appliquer sur un train pour qu'il se déplace tel qu'on l'observe effectivement ? Les forces qui s'appliquent respectent évidemment les bornes définies pour les valeurs de l'accélération. Concrètement, si le train est en retard ou en avance, le conducteur va vouloir corriger cet écart à l'horaire. On va écrire la réponse du conducteur à un écart horaire comme si le train était

soumis à une force. Dans d'autres cas de figure où la sécurité est en jeu, par exemple si le train roule trop vite ou qu'il vient de croiser un signal jaune, alors l'écart à l'horaire devient secondaire pour le conducteur et il doit freiner même s'il est en retard. Ceci peut également se traduire par une force. On définit de cette manière les trois forces déterministes qui sont les suivantes :

1. F_{inter} : l'interaction entre deux trains transmise par les signaux.
2. $F_{v \ max}$: une force qui traduit l'interdiction de dépasser la vitesse limite.
3. $F_{horaire}$: la force horaire qui incite le train à respecter l'horaire.

On remarquera que les forces présentées agissent toutes dans un environnement à une dimension c'est pourquoi à partir de maintenant toutes les forces seront présentées comme des scalaires et non des vecteurs. F_{inter} et $F_{v \ max}$ correspondent à des situations où la sécurité est en jeu. En pratique elles s'appliquent en priorité avant $F_{horaire}$, selon un schéma de non-additivité :

- Si F_{inter} n'est pas nulle, alors $F_{max \ speed}$ et $F_{horaire}$ sont nulles.
- Si $F_{v \ max}$ n'est pas nulle, alors $F_{horaire}$ est nulle.
- Dans tous les cas, il reste au moins l'une de ces trois forces.

On définit une quatrième force : la force aléatoire F_{rand} qui modélise le comportement stochastique de notre système. Cette force représente également la perturbation dont il est question dans la définition de la robustesse vue dans la partie 1.2.4. Dans tous les cas, la résultante des forces qui s'appliquent au train est obtenue en additionnant F_{rand} à la seule force qui s'exprime parmi F_{inter} , $F_{v \ max}$ et $F_{horaire}$. L'indice i représente le numéro d'un train, S_i représente l'état du dernier signal croisé par le train i , Δh l'écart horaire du train, h_d et $t_{friction}$ représentent des temps caractéristiques d'évolution de la force horaire, leur signification est développé en annexe A. Voici les expressions qui décrivent ces quatre forces :

$$F_{inter} = \begin{cases} -a_{max} \cdot m & \text{si } v > v_{secu} \text{ quand } S_i = "jaune" \\ -a_{max} \cdot m & \text{si } v > 0 \text{ quand } d_{i,RS} < l_{secu} \\ 0 & \text{sinon} \end{cases} \quad (2.2)$$

$$F_{v \ max} = -|a_{max}| \cdot m \quad \text{si } v > v_{max}, \quad 0 \text{ sinon} \quad (2.3)$$

$$\begin{aligned} F_{horaire} &= F_{elastiq} + F_{friction} \\ F_{elastiq} &= -m \cdot a_{elastiq} \cdot \tanh(\Delta h / h_d) \\ F_{friction} &= -m \cdot a_{friction} \cdot \tanh(\partial_t(\Delta h) \cdot t_{friction} / h_d) \end{aligned} \quad (2.4)$$

$F_{horaire}$ représente une façon standardisée pour un conducteur de train de réagir à un écart à l'horaire qui peut être positif ou négatif. Comme annoncé précédemment, à partir de maintenant, on appellera les retards des écarts à l'horaire³. Cette force est construite à partir de deux composantes :

- $F_{elastiq}$ analogue à une force de rappel élastique qui s'oppose à un écart horaire.
- $F_{friction}$ analogue à une force de frottement qui s'oppose à une variation de l'écart horaire. Sans cette composante, $F_{elastiq}$ conduirait à osciller sans fin autour de Δh .

$F_{elastiq}$ et $F_{friction}$ peuvent être vues comme les deux termes de la réponse proportionnelle dérivée à un écart horaire Δh . Cette réponse est rendue non linéaire par l'usage de la fonction $tanh$, pour empêcher $F_{horaire}$ de diverger. En effet notre modèle doit rester valide pour un train qui aurait énormément de retard or, pour un tel train, le calcul de la force horaire sans l'utilisation de la fonction $tanh$ conduirait à des forces colossales et irréalistes, et plus le retard serait grand, plus ces forces seraient grandes. En réalité l'accélération d'un train est bornée et par conséquent la force horaire doit être bornée. La fonction $tanh$ permet de tenir compte de cette limitation physique. À partir de cette forme générale, quatre paramètres doivent encore être fixés pour déterminer complètement $F_{horaire}$: $a_{elastiq}$, $a_{friction}$, h_d et $t_{friction}$.

Le calcul détaillé est disponible en annexe A. À l'issue de ce calcul, le temps caractéristique de réponse de la force $F_{horaire}$, t_{car} et la sensibilité de la sous-force $F_{elastiq}$, h_d sont fixés respectivement à 89,35s et à 180s. Ces choix proviennent d'observations expérimentales et d'une discussion avec des experts dans la conduite de trains du réseau Transilien. Il ne s'agit pas cependant de fixer les paramètres pour correspondre au mieux au réseau Transilien, mais bien de construire un modèle assez général. Fixer les valeurs de t_{car} et de h_d nous conduit à donner à $a_{elastiq}$ et à $a_{friction}$ les valeurs respectives de $0.52m.s^{-2}$ et de $0.076m.s^{-2}$.

La notion de force aléatoire a été initialement introduite en physique pour modéliser le mouvement Brownien puis plus tard l'évolution des cours de la bourse. C'est une notion qui s'est répandue jusqu'à aujourd'hui dans de nombreux domaines de la physique pour représenter les phénomènes stochastiques. Dans notre modèle, nous introduisons une force aléatoire F_{rand} qui sert à modéliser des écarts aléatoires aux comportements déterministes définis par les trois autres forces. En créant cette force, on a en tête la variabilité due aux comportements de conduite des conducteurs. Il faut garder à l'esprit que cette force peut tenir compte d'autres sources d'aléas. L'analyse stochastique est puissante en ce qu'elle permet de modéliser des aléas dont on connaît les effets mais sans forcément en connaître les causes.

F_{rand} obéit à un processus de Markov : cela signifie que sa valeur à l'instant $t + dt$ est déterminée uniquement par sa valeur à l'instant t . On notera également que les valeurs prises par une force doivent être bornées. Nous avons choisi de borner cette force avec

3. La convention choisit étant qu'un écart horaire est négatif pour un retard et positif pour une avance.

l'intervalle $[-m \cdot a_{r\ max}, +m \cdot a_{r\ max}]$ ou $a_{r\ max}$ vaut un pourcentage de a_{max} ⁴. En pratique, il est intéressant de bénéficier de la discréétisation numérique de notre système pour définir F_{rand} de la manière suivante⁵ :

- Au premier pas de temps, $t = 0$, F_{rand} est tiré aléatoirement et uniformément sur l'intervalle $[-m \cdot a_{r\ max}, +m \cdot a_{r\ max}]$.
- Pour les pas de temps suivants, la force F_{rand} peut soit garder sa dernière valeur avec la probabilité p soit être tirée aléatoirement et uniformément sur l'intervalle défini précédemment avec la probabilité $1-p$. Connaissant le pas de temps, le temps d'auto-corrélation, t_{corr} peut être ajusté en fixant p .

Les détails de la construction de F_{rand} sont présentés en annexe B. De même qu'avec les paramètres de $F_{horaire}$, le temps d'auto-corrélation de la force F_{rand} , t_{corr} est fixé à 5min. L'amplitude de la force aléatoire s'écrit donc $m \cdot a_{r\ max}$. Pour représenter cette amplitude on utilisera le paramètre sans dimension $a_{r\ max}/a_{max}$. Le choix d'une distribution uniforme résulte d'une volonté de faire un modèle le plus simple possible. Ce choix n'est justifié par aucune observation expérimentale particulière.

2.2.4 Implémentation numérique du modèle

Ce modèle est un système de T trains obéissants à un ensemble de forces que l'on vient de décrire. En appliquant la seconde loi de Newton pour chacun des trains, on voit apparaître un système de T équations différentielles non linéaires et stochastiques du second ordre. Ce système non-déterministe car stochastique est un processus de Markov à représentation d'états continus et à temps continu. À partir d'une condition initiale, le futur de ce système peut être statistiquement déterminé. Mais les définitions par morceaux et les non-linéarités engendrées par la fonction \tanh rendent un calcul analytique quasiment impossible à mener.

Si une étude analytique rigoureuse est difficile, en revanche, il est nettement plus simple de mettre en place une résolution numérique approchée de ces équations différentielles en discréétisant le temps avec un pas temporel constant dans une méthode d'intégration d'Euler⁶. La nature stochastique des équations implique des tirages aléatoires, les résultats ne peuvent alors être interprétés que de manière statistique dans le cadre d'un grand nombre de répétitions de ces résolutions avec renouvellement du tirage aléatoire. C'est la méthode de Monte-Carlo. Nous proposerons dans un second temps des raisonnements analytiques en faisant des hypothèses simplificatrices afin de mieux appréhender le fonctionnement de notre modèle.

Les forces exprimées précédemment sont définies par morceaux. En particulier, l'expression de $F_{v\ max}$ est discontinue lorsque la vitesse limite est dépassée. De même, l'expression de F_{inter} est discontinue lorsque la valeur de la vitesse franchit la vitesse de sé-

4. Ce pourcentage ne dépassera jamais 24% dans ce manuscrit pour que les distances de freinage maximum autorisées par notre infrastructure soient respectées.

5. La discréétisation du système est présentée par la suite.

6. On utilise en pratique les méthodes d'approximation d'une intégrale par des rectangles et des trapèzes avec un pas de temps fixe.

curité ou la vitesse nulle. Discontinuité et discréétisation ne font pas bon ménage : en l'absence de précaution lors de la discréétisation des équations, la valeur de la vitesse va osciller de manière non-amorties autour des valeurs de v_{max} et de v_{secu} . Pire, alors que les équations semblent interdire à juste titre des vitesses négatives, de telles vitesses apparaissent bien dans le système discréétisé. On peut s'affranchir de ces problèmes en cassant artificiellement les discontinuités. De même, si le train roule très lentement, on inhibe toute tentative de la force aléatoire (comme pour les autres forces) qui conduirait à faire reculer le train.

Un pas de temps de $0,1\text{s}$ assure que les erreurs numériques dues au calcul et à la discréétisation durant une résolution numérique restent petites devant les variations dues à la nature stochastique du problème. On observe en effet, entre deux résolutions numériques, des variations dues aux aléas bien supérieures à celles causées par des erreurs numériques. En pratique, nous avons testé différents pas de temps et différentes intensités de la force aléatoire avec la méthode de Monte-Carlo. La comparaison entre ces résultats issus d'une méthode de Monte-Carlo ne peut se faire que de manière statistique, cette comparaison montre que la diminution du pas de temps ne change pas les résultats des simulations.

Le cœur du code de ce modèle est écrit en langage C++, les fichiers de résultats générés par ce code sont ensuite exploités en langage Python.

2.3 Apparition de la congestion

Définition - La congestion

On dit qu'un système ferroviaire est congestionné si au moins un train rencontre un signal orange. Cette définition est une prolongation du concept de congestion utilisée en trafic routier où la congestion se caractérise par une interaction entre un véhicule suivi et un véhicule suiveur qui est forcée de ralentir.

Selon cette définition définie, notre système est congestionné si F_{inter} est non nulle pour au moins un train. Cette définition de la congestion implique pour notre infrastructure qu'au moins un train doit ralentir à la vitesse de sécurité sur au moins un canton.

2.3.1 Organisation des simulations

Dans les conditions initiales, on choisit de mettre tous les trains à équidistance les uns des autres. De plus, on positionne tous les trains à l'heure et à la vitesse prévue par le plan de transport. Les trains ne sont donc ni en retard ni en avance au début de la simulation. Le plan de transport implicite ici correspond à des trains qui restent équidistants au cours du temps.

Pour étudier l'apparition de la congestion, on ne fait varier que trois paramètres : la vitesse maximum autorisée v_{max} , l'amplitude des variations de la force aléatoire qui s'écrit $m \cdot a_{r max}$ et le nombre de trains T. Ces trois paramètres forment un espace à trois dimensions, que les simulations vont couvrir de manière partielle seulement. Considérons le triplet $(v_{max}[km.h^{-1}], a_{r max}/a_{max}[\%], T)$ comme les coordonnées d'un point dans cet espace. L'ensemble des points de cet espace couvert par les simulations est présenté de deux manières en figure 2.3. La particularité de l'ensemble des points parcourus est que le couple $(v_{max}, a_{r max}/a_{max})$ décrit une croix dans le sous-espace à deux dimensions correspondant. C'est une façon de limiter le nombre de simulations sans gêner l'interprétation des résultats.

Dans le cadre de cette étude sur l'apparition de la congestion et pour les besoins de la méthode de Monte-Carlo, les résolutions numériques ont été répétées mille fois pour chacun des points de l'espace des paramètres présenté dans la figure 2.3. Par ailleurs, au cours de cette résolution, les positions des trains sont enregistrées toutes les dix secondes de temps simulé. Comme dit précédemment, la simulation est arrêtée dès l'apparition de la congestion ou après trois heures de temps simulé si la congestion n'est pas apparue. Dans le cas où la congestion apparaît, le temps d'apparition est enregistré.

Enfin, on réalise les simulations avec C = 40 cantons de $l_c = 1000m$, la longueur des trains est quant à elle de $l_t = 200m$. L'ensemble des valeurs utiles pour le modèle est résumé dans le tableau 2.2.

2.3.2 Compétition entre la force aléatoire et la force horaire

La définition de la robustesse donnée au chapitre 1 est trop générale pour pouvoir servir dans ce chapitre et il convient de la préciser. La robustesse représente la stabilité de la performance de notre système. On choisit comme critère de performance pour notre système la non-survenue de la congestion. On mesure la robustesse en mesurant l'amplitude de la perturbation nécessaire pour faire tomber notre système en congestion. Ici la perturbation est la force aléatoire F_{rand} et son amplitude est $m \cdot a_{r max}$. Cette définition restera valable pour la suite de ce manuscrit sauf indication du contraire.

Notre plan de transport qui prévoit que tous les trains sont équidistants et le restent au cours du temps, est le plan de transport le plus robuste qui soit pour un nombre de trains donné. C'est, comme on va le voir un peu plus loin, l'extrême simplicité et homogénéité de l'infrastructure décrite qui permet de trouver le plan de transport le plus robuste de manière triviale. Ce n'est pas évident dans un cas réel.

Dans un premier temps, on s'intéresse uniquement à l'apparition de la congestion et non à son évolution. F_{inter} joue alors un rôle mineur dans le sens où il reste nul pendant toute la durée d'une simulation et que l'on arrête la simulation dès qu'il devient non nul ; signe de l'apparition de la congestion. On peut donc dans le raisonnement qui va suivre oublier l'expression de F_{inter} .

Pour tenter de mieux appréhender le fonctionnement de notre système, on va faire une simplification : on va également oublier momentanément la présence de la force $F_{v \ max}$ pour se concentrer sur la compétition entre les deux forces restantes $F_{horaire}$ et F_{rand} . F_{rand} peut perturber un train à l'heure et l'amener à prendre du retard ou de l'avance, alors $F_{horaire}$ va s'opposer à l'action de F_{rand} . On a vu que F_{rand} pouvait garder la même valeur pendant un certain temps t_{corr} . On se place dans le cadre limité temporellement où cette force prend une valeur non nulle et n'en change pas. À mesure que l'écart horaire créé par F_{rand} croît, la force $F_{horaire}$ croît également. Il se trouve un instant où les deux forces s'équilibrent et où l'écart horaire Δh cesse de croître. La compétition entre ces deux forces conduit à borner l'écart horaire dans la mesure où la force F_{rand} est aussi bornée. La compétition entre ces deux forces est mise en équation dans l'annexe C. On va considérer au regard de ces calculs que les bornes suivantes pour F_{rand} et Δh sont équivalentes de manière approchées⁷ :

$$\begin{aligned} -m \cdot a_{r \ max} &\leq F_{rand} \leq m \cdot a_{r \ max} \\ \Leftrightarrow -h_d \cdot \operatorname{arctanh}\left(\frac{a_{r \ max}}{a_{elastiq}}\right) &\leq \Delta h \leq h_d \cdot \operatorname{arctanh}\left(\frac{a_{r \ max}}{a_{elastiq}}\right) \end{aligned} \quad (2.5)$$

Δh décrit tous les écarts horaires quelle que soit leur nature. Dans le cas où un train est «responsable» de son propre écart horaire comme dans le raisonnement précédent, on appelle cet écart horaire un écart horaire primaire. Ceci est en accord avec la bibliographie présentée au début de ce chapitre. Dans notre modèle, un tel écart horaire peut être positif ou négatif selon que la situation correspond à un retard ou à une avance. Des variations raisonnables des paramètres qu'utilisent ces deux forces nous assurent que cet écart horaire primaire ne dépasse jamais une minute en valeur absolue. Notre modèle étudie donc l'apparition de la congestion dans l'hypothèse où elle est induite par des phénomènes de faible ampleur et qui peuvent rester invisibles dans un système réel⁸. Par exemple, un conducteur qui ne prête pas attention à l'heure et prend un peu de retard rentre dans le cadre de ce modèle. À l'inverse, une avarie sévère entraînant l'arrêt d'un train ne rentre pas dans le cadre de ce modèle.

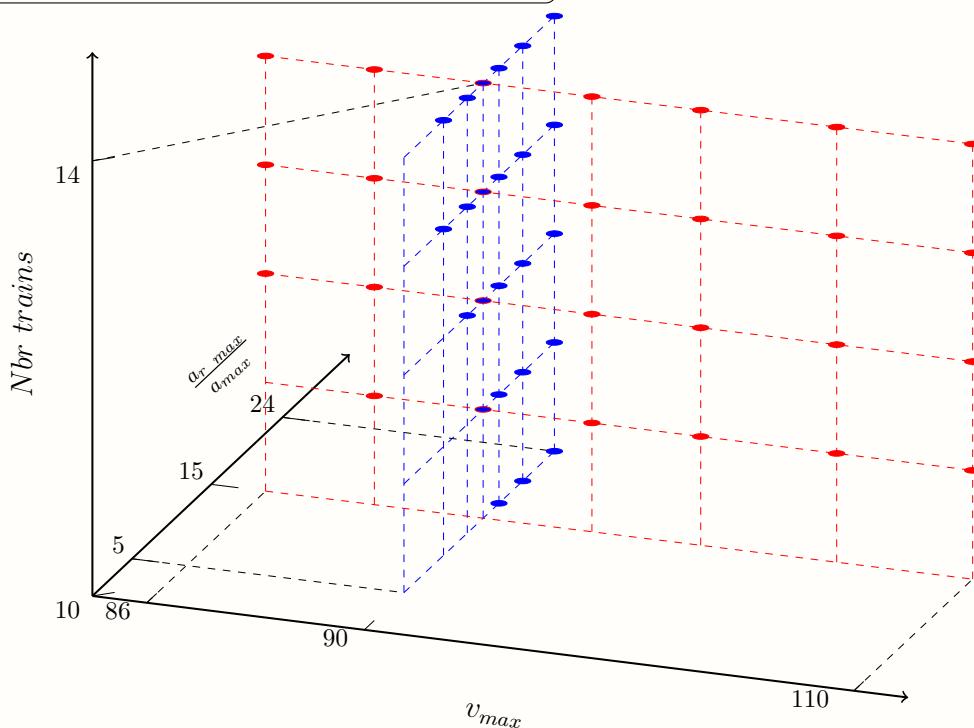
On a affirmé que notre plan de transport était le plus robuste. On peut le démontrer en se reposant sur le raisonnement que nous venons de faire : dans notre plan de transport, tous les sillons sont tracés à égale distance les uns des autres. Une modification de ce plan entraîne nécessairement un rapprochement temporel entre au moins deux sillons. Or, d'après le raisonnement qui précède, à mesure que l'on augmente l'amplitude de la force aléatoire, on augmente aussi l'amplitude de l'écart horaire (cf (2.5)). Tant que cette amplitude reste petite, le système ne tombe pas en congestion. Mais lorsque l'amplitude de l'écart horaire grandit, elle rend la congestion possible et cette congestion va apparaître d'abord là où les sillons ont été tracés proches les uns des autres. Or, par définition de la robustesse de ce système, l'amplitude de l'écart horaire correspond à la taille de la pertur-

7. L'annexe C détaille pourquoi l'équivalence n'est pas totale.

8. selon les moyens d'observation dont on dispose

bation. Pour maximiser la robustesse du plan de transport dans le cadre de notre modèle, il faut donc maximiser la distance minimale entre deux sillons. C'est ce que l'on fait en choisissant de placer les sillons à équidistance les uns des autres.

Figure 2.3 - Espace des paramètres exploré



$$E = \begin{aligned} & (90) \cdot (5, 10, 13, 15, 17, 20, 24) \cdot (10, 11, 12, 13, 14) \\ & + (86, 88, 90, 92, 95, 100, 110) \cdot (15) \cdot (10, 11, 12, 13, 14) \end{aligned}$$

Description graphique de l'espace des paramètres exploré E formant une croix (figure non à l'échelle) et son équivalent écrit sous la forme d'un produit cartésien. Les nœuds de la grille bleue correspondent à tous les points avec $v_{max} = 90 \text{ km.h}^{-1}$ tandis que les nœuds de la grille rouge correspondent à tous les points avec $a_{r \max} / a_{\max} = 15\%$. Les unités sont données par le triplet : $(v_{max} [\text{km.h}^{-1}], a_{r \max} / a_{\max} [\%], T)$. De petits cercles sont ajoutés pour chacun des points où la congestion est apparue dans au moins une simulation parmi les mille menées.

2.3.3 Résultats

Pour chaque point de l'espace des paramètres, on mesure la fréquence d'apparition de la congestion au cours du temps qui est un estimateur statistique de la probabilité d'apparition de la congestion au cours du temps. La figure 2.3 montre tous les points de l'espace des paramètres parcouru et distingue ceux pour lesquels on observe l'apparition de la congestion dans un délai de trois heures. Cette figure ne permet pas de distinguer à quel point la congestion s'est montrée rapide ou lente.

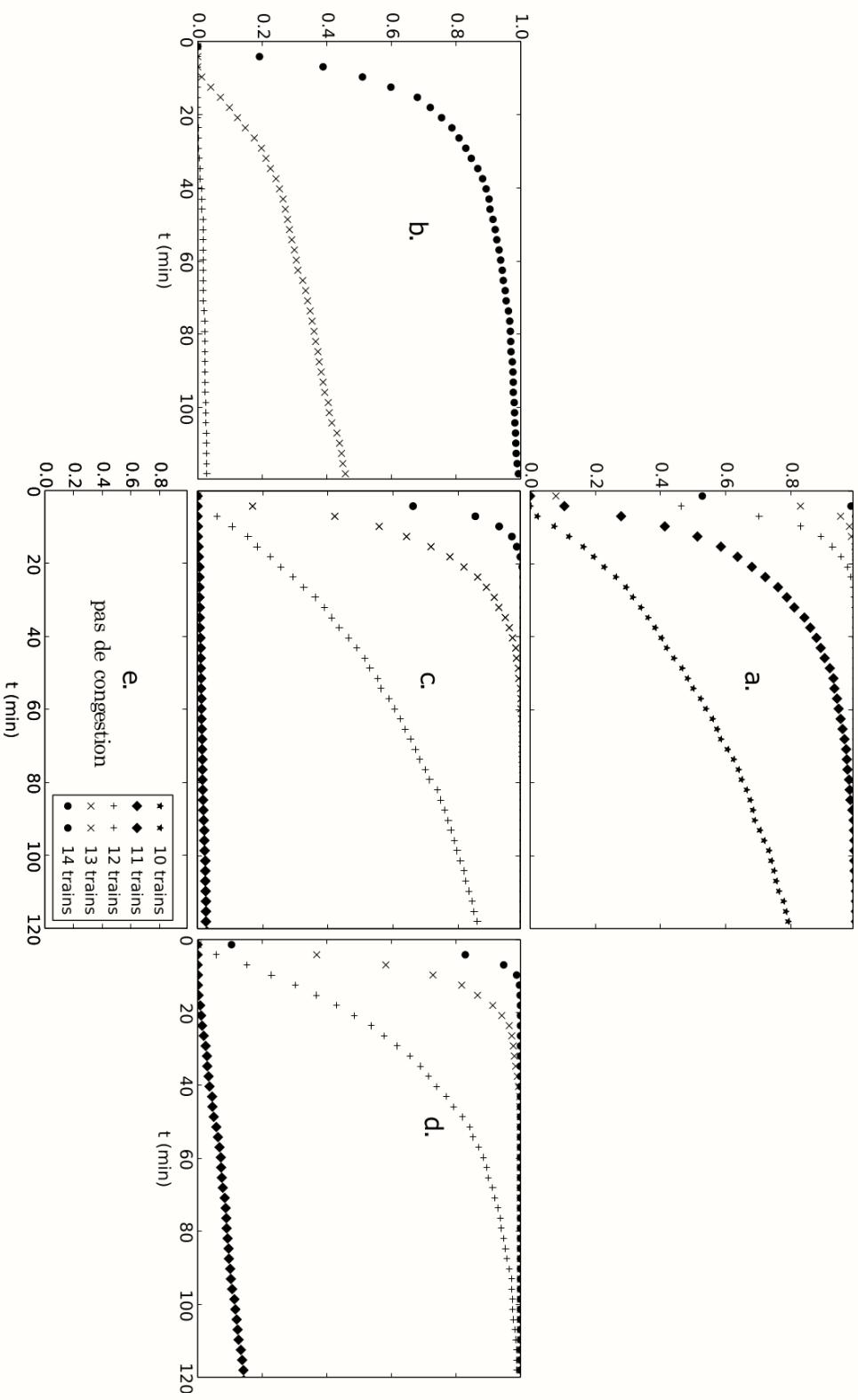
La figure 2.4 représente quelques graphiques d'apparition de la congestion parmi les plus représentatifs. Par rapport à la figure 2.3, ces graphiques représentent les points situés au centre et sur les quatre extrémités de la croix. Comme attendu, on observe que la congestion augmente avec le nombre de trains et l'amplitude de la force aléatoire. Le résultat observé quant à la variation de la vitesse maximum v_{max} est moins intuitif : on observe que la probabilité de tomber en congestion à un instant donné augmente quand v_{max} augmente. Notons aussi que dans tous les cas, la probabilité reste nulle dans les premiers instants. Par exemple sur la figure 2.4.b, on observe qu'avec treize trains aucune congestion n'apparaît dans les dix premières minutes. C'est une observation importante dans les raisonnements qui vont suivre en 2.4.

La figure 2.5 montre des exemples de la distribution des écarts horaires à deux instants : $t = 2,5\text{min}$ et $t = 10\text{min}$. Ces exemples correspondent sur la figure 2.3 au centre et aux quatre extrémités de la croix mais sont représentés pour un nombre de douze trains uniquement. Notons que les systèmes où la congestion est apparue sont de fait exclus de ces distributions puisque les simulations s'arrêtent en cas de congestion. La distribution pour $t = 0\text{min}$ n'est pas représentée, elle correspond aux conditions initiales où tous les trains sont à l'heure cette distribution est donc un pic de Dirac centré en zéro. On observe une dispersion au cours du temps de ces distributions des écarts horaires ; cette dispersion est d'autant plus forte que l'amplitude de la force aléatoire croît. On observe également qu'une vitesse limite trop faible semble limiter les valeurs possibles pour les écarts horaires positifs ; ce phénomène traduit la facilité à prendre du retard comparativement à la difficulté à prendre de l'avance si la vitesse est limitée.

Du fait de la configuration particulière et simplifiée de notre modèle, ainsi que d'un choix de paramètres cohérents avec la réalité mais néanmoins non calibrés sur des cas réels, il faut se retenir de comparer quantitativement les résultats obtenus avec un système réel. Ce n'est d'ailleurs pas la finalité de ce modèle. En revanche, on peut s'intéresser aux mécanismes qui sont mis en avant pour expliquer l'apparition de la congestion et dans c'est dans ce cadre là que l'on présentera au chapitre 4 une comparaison qualitative entre les résultats de notre modèle et une analyse de données du trafic d'un cas réel.

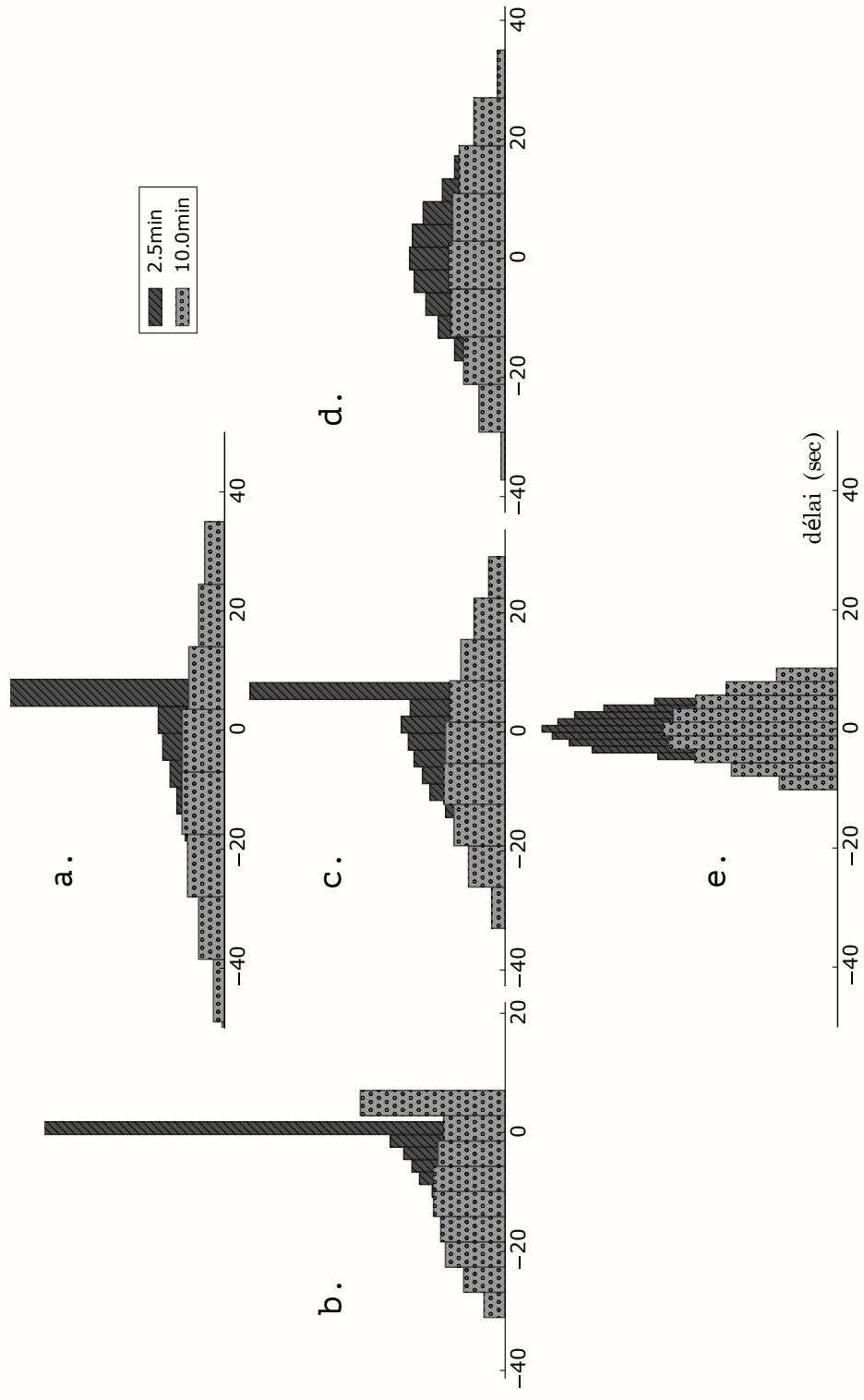
En effet, une partie des courbes présentées, et qui correspondent à des systèmes qui congestionnent avec 99% de chances au bout de quelques minutes, ne sont peut-être pas représentatifs de la réalité ferroviaire, il convient de nuancer ces apparences à la lumière de l'étude sur la quantification de la congestion dans la partie 2.5. Ce travail présente un modèle simplifié pour visualiser comment une force aléatoire, qui modélise entre autres les aléas de conduite des trains, peut les conduire à s'écartez de leurs horaires jusqu'à générer de la congestion. Les trains soumis à cette force aléatoire grignotent les marges temporelles que l'on a disposées entre eux lors de la construction du plan de transport pour augmenter la robustesse. Ce modèle et les résultats que l'on vient de présenter vont maintenant servir à alimenter une analyse de l'apparition de la congestion ferroviaire qui repose sur des processus de Markov.

Figure 2.4 - Probabilité d'apparition de la congestion au cours du temps



Probabilité d'apparition de la congestion au cours du temps pour des systèmes comportant 10 à 14 trains et pour 5 couples différents de (v_{max} , F_{rand}) : **a** : (90,24). **b** : (86,15). **c** : (90,15). **d** : (110,24) **e** : (90,5). Notons que l'apparition de la congestion ne se fait pas immédiatement et que la probabilité reste nulle dans les premiers instants. Lorsque le nombre de trains augmente, le temps pour lequel cette probabilité devient non nulle décroît, de la même manière, le temps nécessaire pour que la probabilité passe de 0 à 1 décroît aussi. La figure e) ne présente aucune courbe car la probabilité d'apparition de la congestion reste nulle quel que soit le nombre de trains, l'amplitude de la force aléatoire ne perturbe pas assez le système pour pouvoir générer de la congestion. Chaque courbe est issue des résultats de 1000 simulations.

Figure 2.5 - Exemples de distributions des écarts horaires primaires



Exemples de distributions des écarts horaires pour 12 trains et 5 couples différents (ν_{max}, F_{rand}) : a) (90,24). b) (86,15). c) (90,15). d) (110,24) e) (90,5). Les conditions initiales (non représentées ici) correspondent à un pic de Dirac centré en 0. Les distributions sont présentées pour des temps de $t = 2.5\text{min}$ et de $t = 10\text{min}$. Ces distributions ont été tracées à partir de 1000 répétitions de la résolution numérique. On observe que ces distributions se dispersent au cours du temps. Cette dispersion semble favorisée par l'augmentation de F_{rand} et de ν_{max} .

Notre modèle est en effet un processus de Markov dont nous n'avons jusqu'ici pas utilisé les propriétés propres à cette famille de processus. On va maintenant s'intéresser à cet aspect de notre modèle dans le cadre de l'apparition de la congestion et s'inspirer de la bibliographie présentée au début de ce chapitre pour tenter de dégager une analyse à la fois de notre modèle mais aussi des résultats qui en sont issus.

2.4 Analyse du modèle et des résultats par des processus de Markov

Dans cette partie, on va expliquer ce qui fait que notre modèle est un processus de Markov puis utiliser les outils adaptés à ces processus afin d'étudier ce modèle dans le cadre de l'apparition de la congestion. Il existe une classe de processus de Markov dont on a déjà parlé dans l'étude bibliographique en 2.1.1, ce sont les distributions de type phase (DTP), elles sont particulièrement adaptées pour une description du fonctionnement de notre modèle mais, on va le voir, sous un angle différent. Le travail suivant nous a été inspiré par les travaux de **BÜKER et SEYBOLD [2012]** et de **MEESTER et MUNS [2007]** et présentés dans l'étude bibliographique.

2.4.1 Processus de Markov

Définition - Processus de Markov

Un processus de Markov est un processus stochastique qui possède une propriété particulière : dans un tel processus, la distribution de probabilité des états futurs d'un système peut être déterminée entièrement si l'on connaît l'état présent du système et sans jamais avoir besoin de connaître le passé de ce système.

Un cours complet sur les processus de Markov peut-être trouvé ici : **BECK et PAUKER [1983]**. Considérons le plus simple des processus de Markov que l'on peut imaginer. C'est un système pouvant se trouver dans deux états. Le premier est un état non-congestionné ou fluide F, le second est un état congestionné C. On peut passer de l'état fluide à l'état congestionné avec un taux de transition s constant. La transition dans l'autre sens est impossible⁹. Ce processus de Markov propose une description de ce qu'il se passe dans notre système d'un point de vue macroscopique. On peut utiliser des processus de Markov pour décrire un système avec une échelle de description plus fine mais ça ne sera pas le cas dans ce manuscrit. Dans le formalisme de la théorie des graphes, ce processus de Markov s'écrit : $F \xrightarrow{s} C$.

⁹. Dans la réalité il est évidemment possible de passer d'un état congestionné à un état fluide, cette hypothèse est en fait spécifique à l'étude de l'apparition de la congestion où on ne s'intéresse pas à l'évolution de cette congestion.

Avec ce processus de Markov, la probabilité d'être dans l'état C au cours du temps est donnée par une distribution exponentielle qui s'écrit : $P_C(t) = f(0) - e^{-s \cdot t}$ avec $f(0)$ qui représente la probabilité d'être dans l'état F à $t = 0$. En utilisant cette distribution exponentielle pour interpoler les courbes qui représentent la probabilité d'apparition de la congestion (voir exemples en figure 2.4), on détermine s . La figure 2.6 montre un exemple d'interpolation de la probabilité d'apparition de la congestion avec la distribution exponentielle (avec 13 trains, une vitesse limite de 95 km.h^{-1} et l'amplitude de la force aléatoire réglée à 15% de $m \cdot a_{max}$). Les points retenus pour les interpolations sont les points pour lesquels la probabilité est inférieure à 0.98, on évite ainsi de pondérer avec les points restant qui ne contiennent presque pas d'informations.

Les interpolations obtenues ne sont pas satisfaisantes pour une raison qualitative : la distribution exponentielle ne peut reproduire ce qui se passe dans notre modèle ferroviaire lors des premières minutes. On voit bien sur la figure 2.4 que la probabilité d'apparition de la congestion reste nulle pendant les quatre premières minutes ce qui n'est pas du tout bien représenté par notre interpolation. En revanche, une seconde interpolation avec une distribution exponentielle calculée en excluant les dix premières minutes montre que la probabilité d'apparition de la congestion converge rapidement et de manière satisfaisante vers une distribution exponentielle.

Cependant il faut garder à l'esprit que le but de cette interpolation n'est pas tant d'approcher les courbes de probabilité d'apparition de la congestion, que de valider un mécanisme explicatif : on cherche à trouver un processus de Markov¹⁰ qui décrive au mieux notre modèle tant dans ses résultats que dans ses mécanismes. Dans cette optique, on ne va pas se satisfaire d'une interpolation par morceaux et on va proposer un processus de Markov un peu plus complexe en augmentant le nombre d'états.

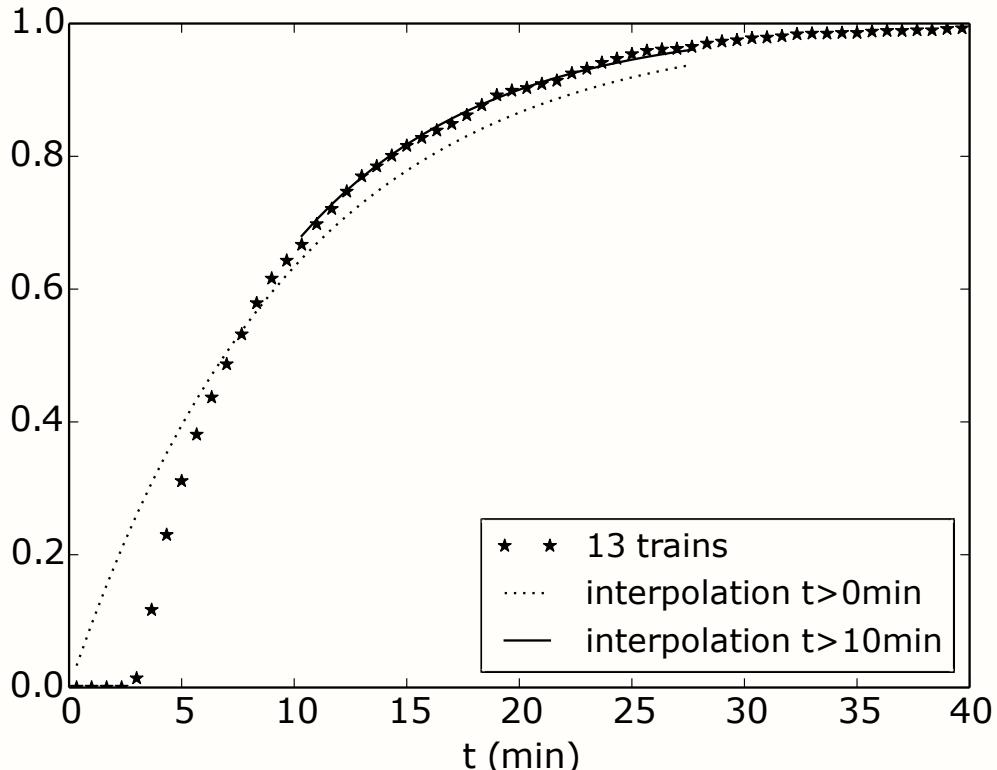
2.4.2 Les distributions de type phase

Les distributions de type phase (DTP) sont des processus de Markov qui décrivent l'évolution d'un système à évolution temporelle continue et avec une description discrète des états que peut prendre le système¹¹. Dans une DTP, il y a au moins un état permanent. Par la suite, on ne va considérer que les DPT avec un seul état permanent. Tous les autres états sont donc transitoires. Une fois qu'un système se retrouve dans un état permanent, alors il n'y a plus de transition possible vers un autre état du système. On peut représenter les DTP par un graphe ou par une matrice.

La figure 2.7 montre ces deux représentations sur un exemple illustratif. Dans le graphe présenté, chacun des nœuds représente un état, les flèches entre les nœuds représentent quant à elles les taux de transition entre les états. Pour un modèle à $m + 1$ états, la matrice correspondante est une matrice carrée de taille $m + 1$ que l'on appelle la matrice des taux

10. on a déjà parlé du caractère markovien de notre modèle en 2.2.3.

11. Il existe aussi des DTP à évolution temporelle discrète ; on ne les utilise pas dans ce manuscrit car malgré la discréétisation temporelle de notre modèle, il reste défini de manière continue temporellement. De plus cela chargerait inutilement les notations.

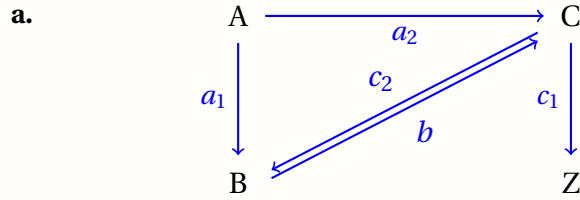
Figure 2.6 - Exemple d'interpolation exponentielle


Exemples d'interpolation exponentielle de la probabilité d'apparition de la congestion. L'interpolation est réalisée une première fois sur l'ensemble des données puis une seconde fois pour les temps supérieurs à $t = 10\text{min}$. La configuration retenue dans cet exemple est : $(v_{max}, a_r max/a_{max}, T) = (95, 15\%, 13)$. On remarquera qu'on n'arrive pas à reproduire la probabilité d'apparition de la congestion lors des premières minutes.

de transition notée Q . La convention est alors de fixer que l'état 0 est l'état permanent alors que les m autres états sont transitoires. $q_{i \neq j}$ représente le taux de transition de l'état i vers l'état j . Les éléments de la diagonale obéissent à une loi de conservation : $q_{i,i} = -\sum_j q_{i \neq j}$. S est une sous matrice carrée de taille m qui ne décrit que les taux de transitions entre les états transitoires. Grâce aux relations qui relient les éléments de la matrice Q entre eux, la donnée de la matrice S suffit à reconstituer de manière univoque la matrice Q . On utilise S pour calculer la probabilité au cours du temps pour le système de se retrouver dans l'état permanent, cette probabilité s'écrit avec une exponentielle matricielle :

$$P(t) = 1 - F e^{S \times t} \cdot \mathbf{1} \quad (2.6)$$

Avec F qui est le vecteur qui décrit l'état initial du système. Ses composantes représentent la probabilité d'être dans chacun des états transitoires à l'instant initial. $\mathbf{1}$ est un vecteur unitaire. Si on considère le système à deux états présenté précédemment, on a S qui se réduit à un simple nombre s , F qui se réduit à un simple nombre : $(1 - c)$ et le vecteur unitaire qui se réduit au nombre 1.

Figure 2.7 - Exemple de représentations des DPT


b.

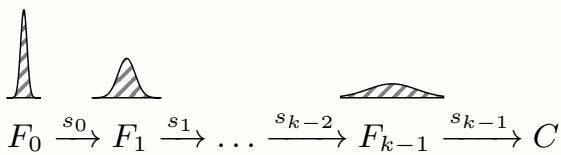
$$Q = \left(\begin{array}{c|cccc} 0 & 0 & 0 & 0 \\ \hline 0 & -(a_1 + a_2) & a_1 & a_2 \\ 0 & 0 & b & -b \\ c_1 & 0 & c_2 & -(c_1 + c_2) \end{array} \right) = \left(\begin{array}{c|c} 0 & 0 \dots 0 \\ \hline S^0 & \mathbf{S} \\ \vdots & \end{array} \right)$$

Exemple générique pour illustrer les deux représentations des DTP. Dans le graphe **a**, Z représente l'état permanent. a_1 , a_2 , b , c_1 et c_2 sont les taux de transition depuis les états A à B, A à C, B à C, C à Z et C à B. La matrice **b** des taux de transition Q où les termes non diagonaux $q_{i \neq j}$ représentent la transition de l'état i vers l'état j . La sous-matrice S se contente de décrire les taux de transition entre les états transitoires.

Robustesse

On a vu au chapitre 1 que la robustesse était définie comme la stabilité de la performance d'un système et dans le chapitre 2 que cette notion pouvait être quantifiée avec la taille de la perturbation nécessaire pour impacter la performance du système. On rappelle que dans notre modèle la taille de la perturbation correspond à l'amplitude de variation de F_{rand} . Elle est donnée par $|F_{rand}| \leq m \times a_{r max}$ (avec m la masse du train). Pour chaque couple constitué par le nombre de trains et la vitesse maximum de l'espace des paramètres (voir figure 2.3), il existe une valeur de l'amplitude des perturbations en dessous de laquelle la congestion n'apparaît pas et au-dessus de laquelle la congestion peut apparaître. On a également vu en 2.3.2 et en annexe C que cette limite en termes de $a_{r max}$ pouvait se traduire en une limite en termes d'écart horaire.

Ce phénomène était déjà illustré sur les figures 2.4 et 2.5. Sur la sous-figure e de la figure 2.5, on a remarqué que la distribution des écarts horaires était peu étalée par rapport aux autres sous-figures. En parallèle le cas correspondant sur la sous-figure e de la figure 2.4 (cette sous figure est vide car la probabilité reste nulle au cours du temps) montre l'absence d'apparition de la congestion. Dans tous les autres cas, la distribution des écarts horaires est plus large et la congestion apparaît. On va maintenant proposer une DTP capable d'une part de mener à une meilleure interpolation des résultats que la distribution exponentielle, d'autre part d'expliquer les conditions nécessaires à l'apparition de la congestion.

Figure 2.8 - Représentation du graphe et de la matrice de notre DTP
a.

b.

$$S = \begin{pmatrix} -s_0 & s_0 & s_1 & 0 \\ & -s_1 & s_1 & \ddots \\ 0 & & \ddots & -s_{(k-1)} \end{pmatrix}$$

Représentation du graphe (figure **a**) et de la matrice (figure **b**) S de notre DTP. Les F_i sont les k états non-congestionnés ; on les a représentés liés à une distribution des écarts horaires (voir la figure 2.5). La largeur de ces distributions augmente avec i . C est l'état congestionné. Dans cette figure les distributions des écarts horaires représentées sont purement illustratives, seule la variation de leur largeur est importante pour le raisonnement. Enfin on remarque que si le taux de transition entre les états F_{k-1} et C , qui vaut s_{k-1} est nul alors il n'y a pas de congestion.

Notre proposition de distribution de type phase

Toutes ces observations nous conduisent à suggérer la DTP qui est présentée dans la figure 2.8. Les F_i sont les k états non-congestionnés et C représente l'état congestionné. L'état initial de ce processus de Markov est ici F_0 et il correspond à un pic de Dirac en termes de distribution des écarts horaires. Ce sont, dans ce processus de Markov, les transitions successives depuis l'état F_0 à travers F_1, F_2, \dots qui font que la distribution des écarts horaires s'élargit au cours du temps ainsi qu'on l'a observé dans la figure 2.5. Enfin, l'apparition de la congestion se fait au moment où la distribution des écarts horaires est devenue suffisamment large et correspond à un système situé dans le dernier état non-congestionné F_{k-1} avant l'état congestionné.

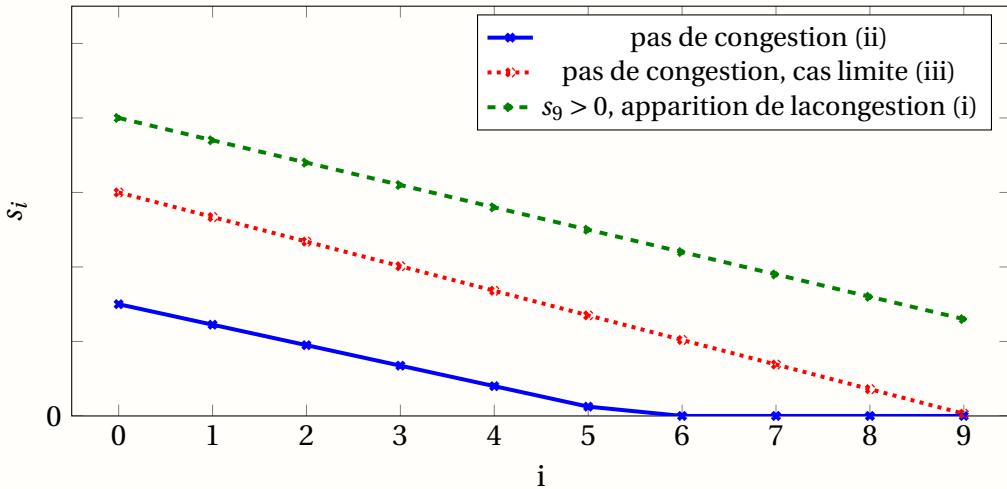
Cependant le modèle reste encore trop général et les valeurs de tous les s_i restent à déterminer. On va diminuer le nombre de variables à la lumière des observations que l'on peut faire sur notre modèle : dans la sous-figure e de la figure 2.4, le système ne congestione pas. Cela implique que dans cette configuration on a au moins : $s_{k-1} = 0$. Le même sous-cas e de la figure 2.5 montre un élargissement de la distribution des écarts horaires au cours du temps ce qui implique que dans cette configuration on a au moins : $s_0 > 0$.

À partir de ces observations et de ce raisonnement, on fait l'hypothèse de la relation convenable la plus simple possible entre les S_i : $s_i = s_0 + i \cdot \Delta s$ avec $\Delta s < 0$. On remarque que ce modèle pourrait conduire à des situations où l'un des coefficients s_i serait négatif ce qui serait absurde. On propose donc une petite correction à la formule précédente :

$$s_i = \max(0, s_0 + i \cdot \Delta s) \quad \text{avec} \quad \Delta s < 0 \quad (2.7)$$

Cette équation et les différents cas de figure qu'elle engendre sont illustrés par la figure 2.9. Essayons maintenant d'interpoler nos courbes de probabilité d'apparition de la congestion avec cette DTP en fixant $k = 10$ et en laissant libres les deux paramètres s_0 et Δs .

Figure 2.9 - Représentation de $s_i = f(i)$ pour différents cas de figure



Représentation de $s_i = f(i)$ dans trois cas de figure. $s_i = \max(0, s_0 + i \cdot \Delta s)$ avec $\Delta s < 0$. Dans cette figure $k = 10$. Les trois cas de figures représentés sont : (i) le taux de transition de F₉ à C, vaut s_9 et est strictement positif, la congestion apparaît. (ii) quand plusieurs s_i sont nuls dont au moins s_9 , alors la congestion est impossible. (iii) Quand $s_9 = 0$ mais que $\forall i \neq 9, s_i > 0$ alors on est dans un cas limite situé entre les deux premiers cas présentés. La congestion n'apparaît pas.

Interpolation

Pour l'interpolation avec les courbes de probabilité d'apparition de la congestion, on utilise la méthode non-linéaire des moindres carrés (implémentée en python dans la bibliothèque scipy et dans la fonction curve_fit). Ainsi on ne détermine pas entièrement la forme que l'on veut voir pour les s_i et on laisse le programme trouver les paramètres qui permettent selon la norme des moindres carrés de réduire au maximum les écarts entre les résultats de nos simulations et la fonction d'interpolation. On répète cette méthode pour des valeurs de k variant entre 2 et 12, k étant le nombre d'états non-congestionnés. On constate à l'aide de l'erreur quadratique moyenne (EQM) que les plus grandes valeurs de k conduisent à une meilleure précision.

La figure 2.10 montre un exemple (correspondant au point $(v_{max}, a_r max / a_{max}, T) = (95, 15\%, 13)$ de l'espace des paramètres) de l'évolution de l'erreur quadratique moyenne en fonction de k . On choisit de fixer $k = 10$ une fois pour toute. C'est un bon compromis entre précision et rapidité de calcul sachant qu'avec l'augmentation de k , on doit gérer des

matrices de plus en plus grosses. À titre d'exemple sur cette figure $MSE(k = 10) = .0026$ et $MSE(k = 11) = .0025$.

La figure 2.11 montre ce que donne un exemple d'interpolation. On utilise la même courbe de probabilité d'apparition de la congestion que dans la figure 2.6. L'interpolation qui en résulte est de meilleure qualité qu'avec la distribution exponentielle, en particulier on arrive à reproduire partiellement le fait que dans les premières minutes la probabilité d'apparition de la congestion reste nulle. Toutefois, les résultats d'interpolations sont mauvais quand la vitesse limite est trop proche de la vitesse planifiée, en pratique quand $v_{max} = 86 km.h^{-1}$. Nous n'avons pas trouvé d'explication simple de ce fait, notons toutefois que cette vitesse maximum est peu réaliste, elle correspond à une marge de vitesse de seulement $1 km.h^{-1}$. Le système simulé est ainsi sans doute peu réaliste.

Les raisons de ces mauvais résultats d'interpolation peuvent se comprendre en étudiant le fonctionnement de notre modèle dans le cas limite où la marge de vitesse est nulle. Dans ce cas, tout retard pris ne peut être rattrapé. Au début de la simulation, les trains sont à l'heure. Il est alors possible selon les aléas de la force aléatoire qu'un train prenne du retard et déclenche de la congestion si le train qui le suit prend un retard moindre. En revanche, après une période transitoire, si la congestion n'a pas déjà eu lieu, les retards s'équilibrent en prenant la valeur du retard maximum possible, alors l'écart temporel entre les trains redevient le même qu'au début de la simulation. Or, pour que la congestion apparaisse, il faut des écarts variables entre les trains. Une telle situation n'est pas réaliste du point de vue de notre formulation PTD. Il est donc normal que lorsque l'on s'approche de cette configuration, cette formulation PTD donne de mauvaises interpolations. Mais heureusement cette situation ne peut pas non plus correspondre à une situation réelle. Ces mauvais résultats ne sont donc pas déterminants.

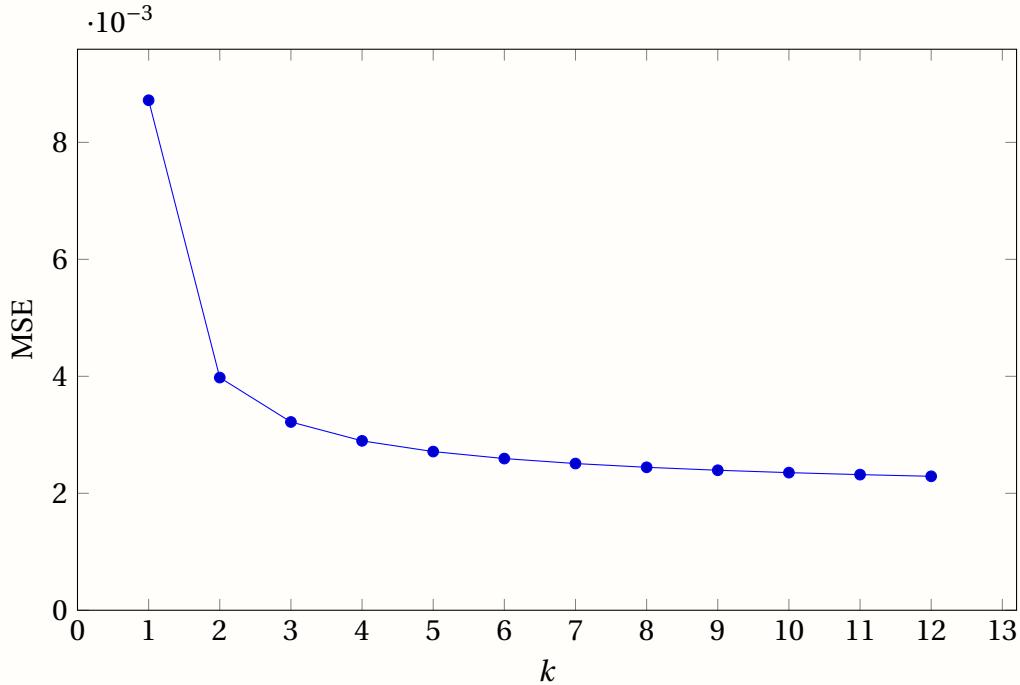
2.4.3 Lien avec le modèle

Cette DTP est, comme on vient de le voir, définie par deux paramètres : s_0 et Δs . Ces paramètres ne semblent cependant pas avoir de lien avec le modèle développé. On peut conduire une analyse mathématique pour choisir un couple de paramètres plus pertinent :

- Une écriture en série de Taylor de $P(t) = 1 - F.e^{S \times t} \cdot \mathbf{1}$ autour de $t = 0$ montre que le paramètre dominant aux temps courts est le produit des s_i noté $\prod s_i$.
- L'analyse de $\lim_{t \rightarrow \infty} (P(t))$ montre que $s_{k-1} = s_9$ domine aux temps longs¹².

Avec ce nouveau couple de paramètres ($\prod s_i$, s_9), c'est s_9 qui est de loin le plus sensible aux variations des paramètres de notre modèle et $\prod s_i$ est mis de côté car il ne représentait pas des variations significatives lorsque les paramètres de modèle variaient. La figure 2.12 montre s_9 en fonction du nombre de trains, de la vitesse maximum et de l'amplitude de la force aléatoire. Par construction, s_i est décroissant en i (voir équation (2.7)). s_9 est donc la plus petite des valeurs des s_i et joue un rôle important : si $s_9 = 0$, alors il n'y a jamais de transition de l'état F_9 à l'état C et la congestion ne peut donc pas apparaître. Au contraire, si

12. s_9 correspond dans le cas où $k = 10$ à la transition de l'état F_9 vers l'état C.

Figure 2.10 - Erreur quadratique moyenne en fonction de k


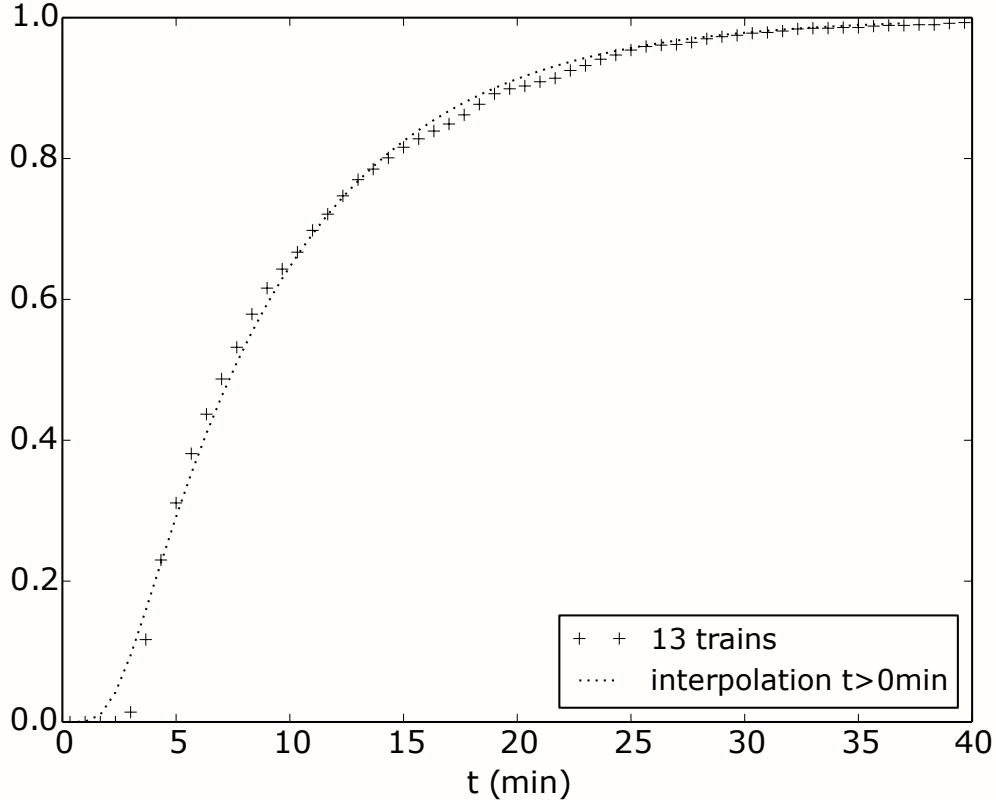
Exemple du calcul de l'erreur quadratique moyenne en fonction de k correspondant au point $(v_{max}, a_{r\ max}/a_{max}, T) = (95, 15\%, 13)$ de l'espace des paramètres. $k = 1$ correspond à une interpolation avec la distribution exponentielle. Lorsque k grandit, l'erreur devient plus petite mais dans le même temps, le temps de calcul augmente. On retient $k = 10$ comme compromis entre erreur et temps de calcul pour la suite de ce manuscrit. Il faut noter que quel que soit k , le processus de Markov DTP ne contient que deux paramètres, s_0 et Δs .

s_9 devient très grand, alors tous les autres s_i deviennent grands aussi et la DTP montre une congestion quasi-instantanée. (voir figure 2.9). Il y a un moyen de déterminer ces deux limites de comportements et de les expliciter en fonction des paramètres utilisés dans notre modèle (voir figure 2.12).

Cas limite : $s_9 \rightarrow 0$

On se place dans le cadre de notre modèle, on considère les positions de deux trains consécutifs et on cherche à expliciter les situations où la congestion peut apparaître. Si $F_{rand} = 0$ la distance de la tête du premier train à la tête du second s'écrit $d_0 = l_C \times C/T$. La configuration de F_{rand} qui rapproche le plus les trains les uns des autres est la suivante : $F_{rand} = -m \times a_{r\ max}$ pour le train suivi et $F_{rand} = m \times a_{r\ max}$ pour le train suiveur. En considérant, comme c'est fait dans l'annexe C et dans la figure 2.13, l'équilibre entre la force aléatoire et la force horaire on trouve :

$$F_{rand} = F_{horaire} \Leftrightarrow \Delta x = \Delta h \times v_{plan} = v_{plan} \times h_d \times \tanh^{-1}(a_{r\ max}/a_{elastiq})$$

Figure 2.11 - Exemple d'interpolation DTP


Exemple d'interpolation DTP avec $k = 10$ dans notre formulation de la DPT. le triplet de paramètres $(v_{max}, a_{r max}/a_{max}, number of trains)$ vaut $(95, 15\%, 13)$.

Dans ce cas, la distance entre la tête des deux trains devient :

$$d = d_0 - 2 * \Delta x = l_C \times C/T - 2 \cdot v_{plan} \times h_d \times \tanh^{-1}(a_{r max}/a_{elastiq})$$

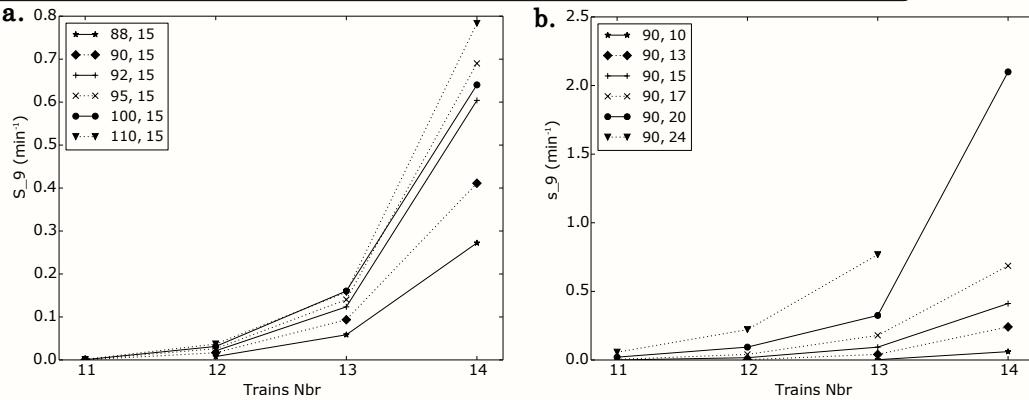
Si $d = d_0 - 2 \cdot \Delta x > 2 \cdot l_C + l_T$ Alors les deux trains sont trop éloignés pour interagir et la congestion n'apparaît pas (voir figure 2.13). La limite $s_9 \rightarrow 0$ s'exprime dans l'égalité suivante :

$$l_C \times (C/T - 2) - l_T = 2 \cdot \Delta x = 2 \cdot v_{plan} \times h_d \times \tanh^{-1}(a_{r max}/a_{elastiq})$$

Et le nombre de trains en dessous duquel la congestion ne peut apparaître et s'écrit :

$$T = l_C \times C / (l_T + 2l_C + 2 \cdot v_{sched} \times \tanh^{-1}(a_{r max}/a_{elastiq})) \quad (2.8)$$

La figure 2.14 illustre cette limite en dessous de laquelle la congestion est impossible et la superpose avec l'espace des paramètres explorés, et en mettant en valeur les points où la congestion est apparue. On remarque que cette limite n'est pas strictement respectée, cela est a priori dû aux approximations faites dans l'annexe C. En effet, pour le point qui correspond au triplet des valeurs suivantes ($v_{max} = 90 \text{ km.h}^{-1}$, $a_{r max}/a_{max} = 17\%$, $T = 10$) la congestion apparaît alors que selon le raisonnement précédent elle ne devrait pas apparaître. Il faut cependant relativiser cette apparition de la congestion qui n'a eu lieu que pour trois simulations sur un total de mille réalisées.

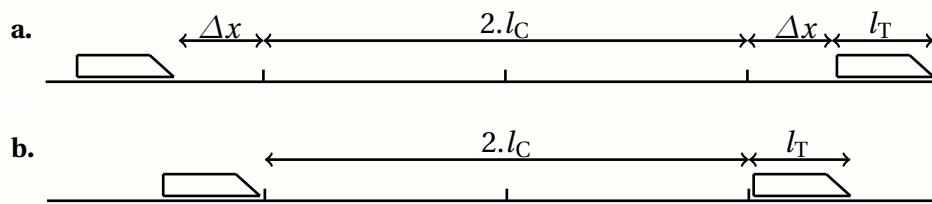
Figure 2.12 - Évolution de s_9 en fonction des paramètres du modèle


Évolution de s_9 en fonction des paramètres du modèle. La légende présente le couple ($v_{max} [\text{km.h}^{-1}]$, $\%F_{rand} [\%]$) associé à chacune des courbes. L'échelle n'est pas la même sur les deux graphes. Le graphe **a** représente la variation de s_9 en fonction de la vitesse maximale et du nombre de trains, $\%F_{rand}$ reste alors constant et vaut 15%. Le graphe **b** représente les variations de s_9 en fonction de l'amplitude de la force aléatoire et du nombre de trains, v_{max} reste alors constant et vaut 90 km.h^{-1} .

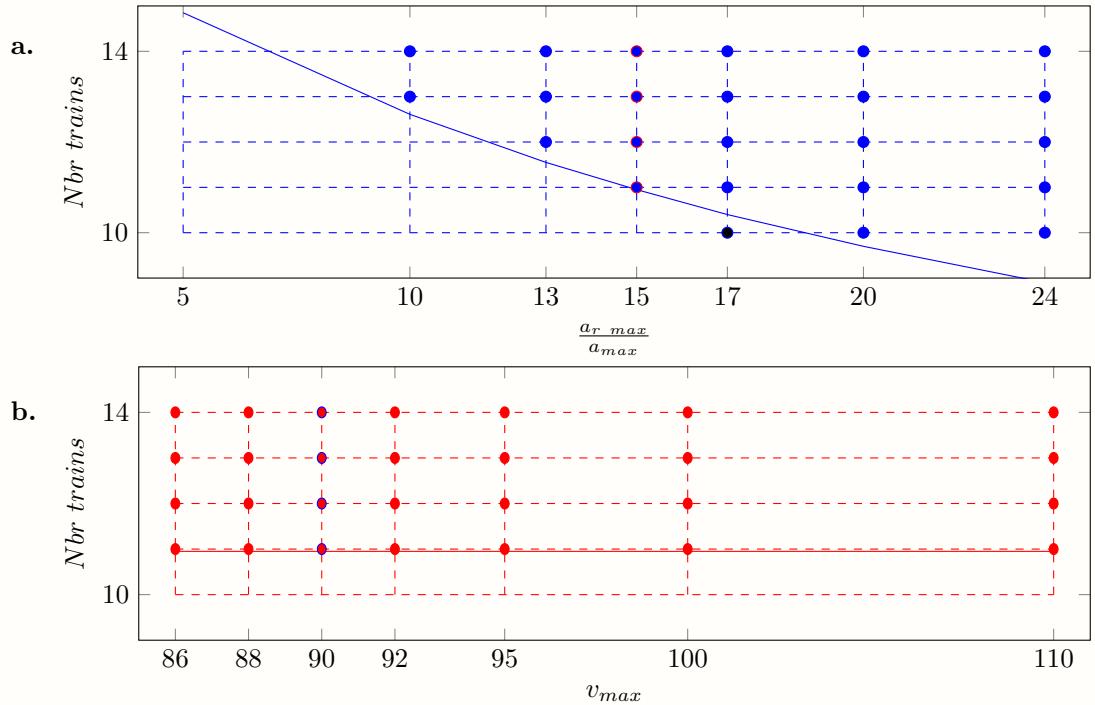
La limite $s_9 \rightarrow \infty$

Cette limite est atteinte lorsque la congestion apparaît même pour de très faibles perturbations. En dessous de cette limite, la congestion apparaît même sans perturbation. Cela veut dire qu'on ne peut concevoir un plan de transport en dessous de cette limite. La figure 2.13 présente aussi cette limite qui correspond à l'inégalité :

$$d_0 > 2.l_C + l_T \Leftrightarrow l_C \times (C/T - 2) - l_T > 0 \quad (2.9)$$

Figure 2.13 - Illustration des distances limites entre deux trains


Distance entre deux trains illustrant deux cas limites. La voie représentée est divisée en cantons. La sous-figure **a** représente la distance limite entre deux trains au-dessus de laquelle la congestion est impossible. C'est la limite $s_9 \rightarrow 0$. Δx représente en distance l'avance ou le retard maximum induit par la force aléatoire. La sous-figure **b** représente la distance limite en dessous de laquelle la congestion apparaît même sans perturbation. C'est la limite $s_9 \rightarrow \infty$.

Figure 2.14 - Limite d'apparition de la congestion


Comparaison entre l'approche analytique et l'approche simulée de la limite d'apparition de la congestion avec en fond l'espace des paramètres explorés (voir figure 2.3). La grille bleue de la sous-figure **a** représente les cas où $v_{\max} = 90 \text{ km.h}^{-1}$. La grille rouge de la sous-figure **b** représente les cas où $a_r \max / a_{\max} = 15\%$. Les unités du triplé des paramètres sont ($v_{\max} [\text{km.h}^{-1}]$, $a_r \max / a_{\max} [\%]$, T). Des cercles ont été tracés pour les points de l'espace des paramètres où la congestion est apparue au moins une fois dans les simulations. Les courbes représentent la limite en dessous de laquelle la congestion ne peut pas apparaître. Ces courbes sont évaluées de manière analytique mais avec certaines approximations, ce qui pourrait expliquer pourquoi cette limite n'est pas parfaitement respectée (sous figure **a** point en (90, 17%, 10)). Il faut cependant relativiser l'importance de l'apparition de la congestion pour ce point car seul 0,3% des simulations congestionnent. L'équation de ces courbes est la formule (2.9).

2.5 Quantification de la congestion

On a jusqu'ici étudié l'apparition de la congestion sans qu'il soit question de sa quantification. Nous utilisons toujours le modèle développé en 2.2 pour la quantifier en laissant de coté la formulation DTP présentée en 2.4.2. En effet, elle ne contient qu'un seul état congestionné ce qui est contraire à la notion de quantification. Il est en effet légitime de se demander, une fois que l'on sait le système congestionné, si la qualité de service offert par les trains reste proche de la qualité de service prévue par le plan de transport ou si au contraire cette qualité est fortement dégradée. Nous allons voir qu'il est possible de prendre à la lumière des résultats qui vont suivre une définition différente de la performance de notre système.

Nous allons faire évoluer en deux versions l'indicateur de performance de notre système. Dans un premier temps nous allons nous intéresser à la vitesse moyenne instantanée de tous les trains du système et dans un second temps nous allons utiliser ces vitesses pour construire un diagramme fondamental de ligne ferroviaire DFLF. Ce diagramme qui est inspiré des DF que l'on a introduits en 1.3 sera présenté à cette occasion. Le dernier chapitre de ce manuscrit sera l'occasion de revenir sur ce DFLF plus en détails. La mesure des vitesses moyennes et la construction de ce diagramme reposent toujours sur le modèle présenté en 2.2. Dans ce but, le programme l'implémentation numérique du modèle a été légèrement modifié, nous allons maintenant présenter ces quelques différences.

2.5.1 Organisation des simulations

Tout d'abord en ce qui concerne les conditions initiales des simulations, les paramètres physiques du système ainsi que l'espace des paramètres étudiés, il n'y a aucun changement. Les simulations s'arrêtent désormais après une heure de temps simulé et les variables récoltées ne sont plus les mêmes. Toutes les vingt secondes de temps simulé¹³, le programme calcule la vitesse moyenne instantanée sur l'ensemble des trains du système. On note cette vitesse \bar{v}^T , écriture qui permet de distinguer que la vitesse a été moyennée sur l'ensemble des trains T. Il est important de signaler que dans un contexte de manque de temps le nombre de simulations dans la méthode de Monte-Carlo a été ramené à seulement cinq cents.

2.5.2 Dynamique de la vitesse instantanée moyenné sur les trains

Les données sont ensuite analysées grâce à un programme python : à partir des vitesses moyennes instantanées sur tous les trains de chaque simulation, on calcule à nouveau une moyenne sur l'ensemble des 500 simulations. On note cette vitesse $\bar{v}^{T,500}$ ce qui permet de voir que l'on a moyenné d'abord sur l'ensemble des trains et ensuite sur l'ensemble des simulations. La différence entre cette vitesse et la vitesse planifiée est un donc indicateur

13. Dans le cadre de l'étude sur l'apparition de la congestion, on récoltait les positions de tous les trains. Les variations de comportement individuel d'un train impliquaient de récolter les postions toutes les dix secondes. À présent nous nous intéressons à la vitesse moyenne instantanée des trains qui, du fait de son caractère moyen, présente une évolution plus lente. Relever ces vitesses toutes les vingt secondes est pertinent.

de quantification de la congestion. On calcule également le premier et le dernier quartile de la distribution de \bar{v}^T en fonction du numéro de la simulation. Quelques graphiques les plus significatifs présentent ces résultats dans la figure 2.15.

On remarque que la congestion n'a pas la même gravité dans tous les cas. On peut également remarquer que la congestion est d'autant plus grave que la probabilité d'apparition de la congestion croissait rapidement (voir figure 2.4 pour comparaison). De plus on remarque que la congestion est parfois de faible gravité au point que la vitesse moyenne des trains $\bar{v}^{T,500}$ vaut toujours la vitesse planifiée. Dans ces cas-là l'indicateur que l'on a construit en définissant différemment l'objectif de performance représente une amélioration notable mais n'est toujours pas parfait : il est évident que de petits retards locaux sont observés et que la qualité du service est légèrement dégradée même si ces retards sont rattrapés par la suite.

On distingue sur ces courbes, un régime transitoire et un régime permanent. Les 500 simulations, à l'instant $t = 0\text{ min}$, se trouvent dans la même situation, à savoir une situation non congestionnée. Au cours du temps, les aléas contribuent à ce que les vitesses moyennes des trains dans les différentes simulations divergent les unes des autres. Les aléas contribuent aussi à l'apparition de la congestion. Le régime transitoire correspond au passage entre le système non congestionné à $t = 0$ et le système congestionné un peu plus tard. Lorsque le système est congestionné, certains trains continuent de rouler à voie libre alors que d'autres sont forcés de ralentir. À la dispersion des vitesses due aux aléas, il faut donc ajouter une dispersion des vitesses due à la congestion. À mesure que l'on augmente l'amplitude des aléas ou le nombre de trains la proportion des trains qui roulent à voie libre, en régime permanent, diminue, la vitesse moyenne $\bar{v}^{T,500}$ diminue donc aussi.

La figure 2.16 reprend la figure 2.14 en ajoutant une information sur la quantification de la congestion : on représente pour chaque point de l'espace des paramètres pour lequel la congestion apparaît, la chute de vitesse (en km.h^{-1}) par rapport à la vitesse planifiée. Une chute de 0 km.h^{-1} indique que si la congestion a bien lieu, les trains sont capables de compenser entièrement les chutes de vitesse qu'ils subissent sur quelques cantons en roulant plus vite sur les autres cantons. Grâce à cet indicateur on peut désormais distinguer trois états de trafic :

- L'état fluide : la congestion n'apparaît pas.
- L'état congestionné résilient : la congestion apparaît mais n'entraîne pas une chute de la vitesse moyenne des trains. Des petits retards apparaissent mais chaque train qui prend du retard est ensuite capable de le rattraper.
- L'état congestionné dégradé : la congestion apparaît, une chute de la vitesse moyenne des trains est constaté, les retards s'accumulent.

Dans l'état du trafic de congestion résiliente, chaque train vit la situation suivante : un train roule à voie libre, rencontre un signal fermé, il ralentit, prend du retard, rencontre à nouveau un signal à voie libre, et rattrape son retard sur les cantons suivants en utilisant la marge de vitesse. Il convient de remarquer que si le système se trouve dans cet état, alors notre indicateur de quantification de la congestion indique une perte de vitesse de 0 km.h^{-1} . En réalité, notre système subit comme on vient de le dire de petits retards locaux

qui peuvent localement la ponctualité et donc la qualité du service offert aux voyageurs. De plus notre modèle ne permet pas de le voir mais dans un système ferroviaire réel, plus complexe, ces petits retards peuvent entraîner d'une part des problèmes de synchronisation des flux ferroviaires sur les convergents et d'autre part des retards plus grands pour les passagers qui sont susceptibles de manquer des correspondances. Il convient d'être bien conscient de ces défauts surtout que le second indicateur de la congestion, le diagramme fondamental de ligne ferroviaire que nous allons maintenant présenter, ne pallie pas ces défauts non plus. Il permettra en revanche de mettre en évidence dans le chapitre 4, l'existence d'un état de congestion dégradée sur des morceaux de deux lignes du réseau Transilien, les lignes B et C.

Chez Transilien la régularité d'un train était historiquement calculée entre son terminus et son origine. Selon cet indicateur, une situation de congestion résiliente qui se développerait au milieu de la ligne n'aurait pas d'impact sur la régularité des trains, mais pourrait en revanche localement impacter des passagers. Cet indicateur a été remplacé en 2012 par un indicateur de ponctualité voyageur calculé par le logiciel VENDOME. En croisant la base de donnée qui donne les retards d'un train dans chaque gare et en pondérant ces retards par une estimation de la quantité de passagers qui descendent dans cette gare, on obtient un indicateur plus proche du ressenti des clients.

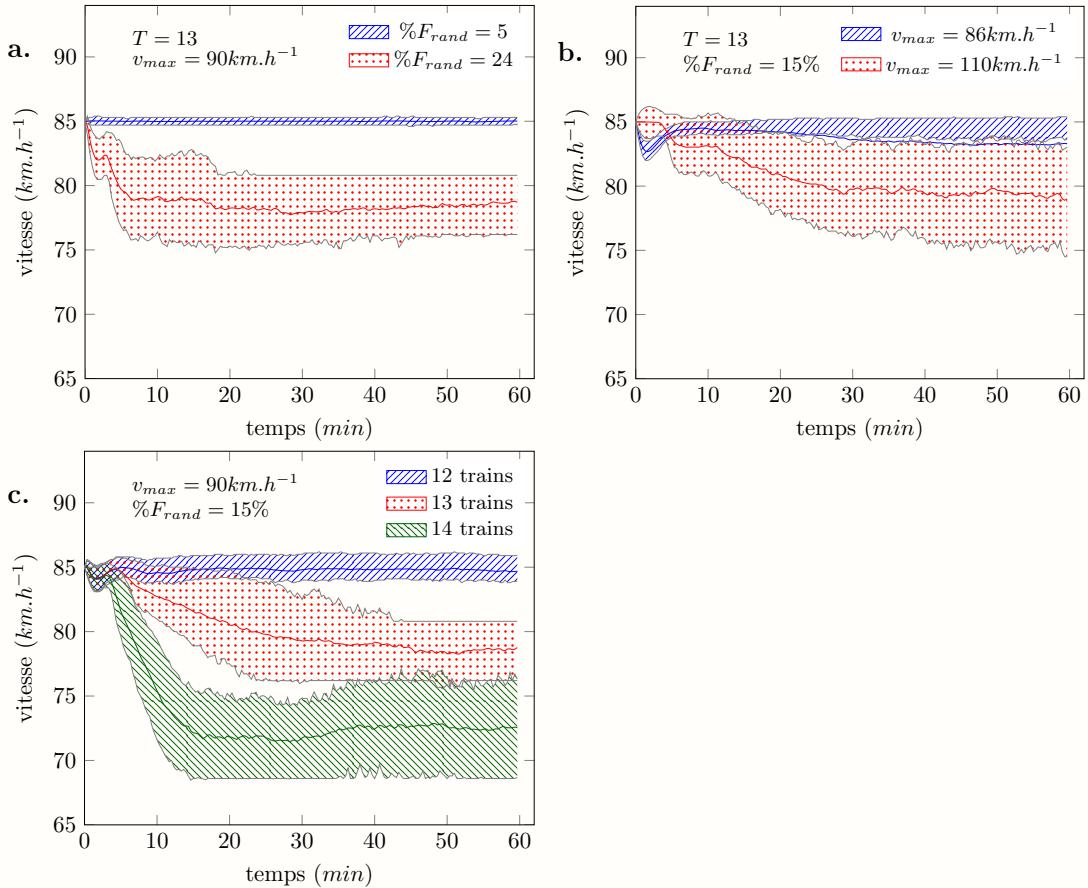
2.5.3 Diagramme Fondamental de Ligne Ferroviaire

Le Diagramme Fondamental de Ligne Ferroviaire (DFLF) représente le débit moyen d'une ligne ou d'une portion de ligne ferroviaire en fonction de la concentration de trains sur cette même ligne ou portion. Les débits calculés sont donc moyennés sur la longueur de la portion de ligne mais aussi sur une fenêtre de temps. La figure 2.17 représente un DFLF pour lequel les points ont été obtenus de la manière suivante :

- Pour $v_{max} = 90 \text{ km.h}^{-1}$ et pour une intensité de la force aléatoire et du nombre de trains variant respectivement entre 5% et 24% d'une part et entre 1 et 14 d'autre part, on a répété les simulations 500 fois dans le cadre de la méthode de Monte-Carlo.
- Pour chaque simulation le débit moyen est obtenu sur l'ensemble de la ligne simulée et sur la fenêtre temporelle des trente dernières minutes des simulations après stabilisation de $\bar{v}^{T,500}(t)$.
- Pour chaque valeur du nombre de trains on obtient une concentration et une distribution des débits moyens, la moyenne de cette distribution et son écart type sont représentés sur la figure 2.17.

La fenêtre spatio-temporelle choisie tend à gommer les aspects dynamiques que pourrait présenter le DFLF. En effet, la prise en compte de l'ensemble de notre ligne ne permet pas de faire fluctuer le nombre de trains au cours d'une simulation. De même la fenêtre temporelle est choisie suffisamment large et elle exclut le régime transitoire des premiers instants que l'on peut voir dans la figure 2.15.

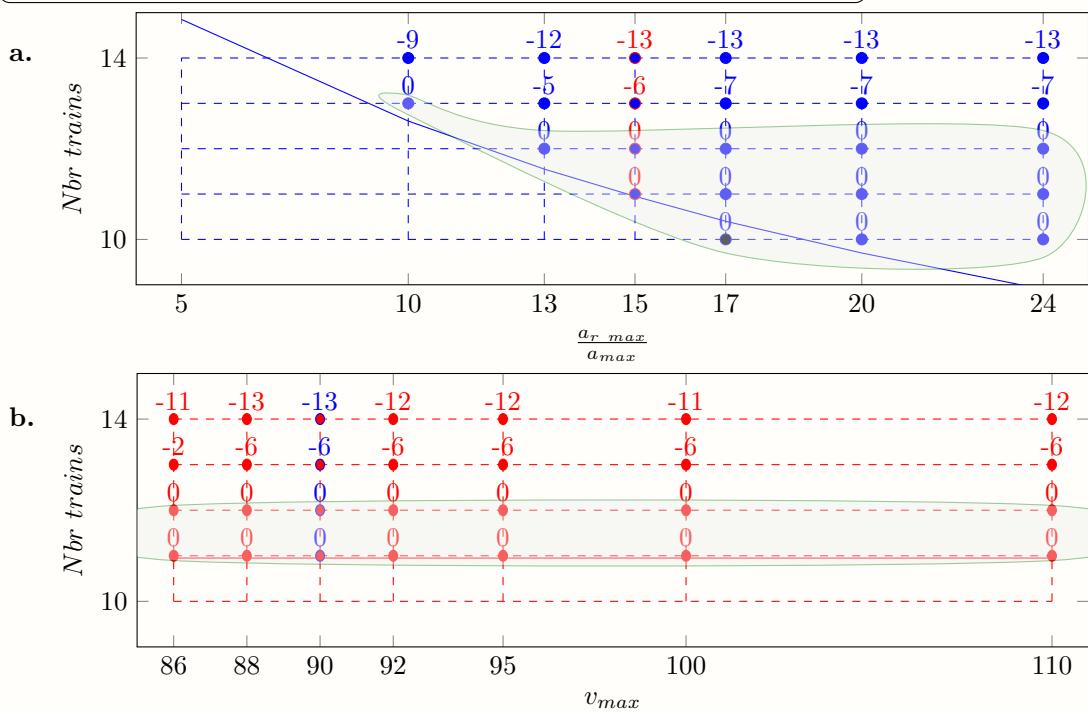
Tant que les retards ne s'accumulent pas, c'est-à-dire tant que le système ne passe pas dans un état congestionné dégradé, le plan de transport initialement prévu peut être

Figure 2.15 - $\bar{v}^{T,500}(t)$


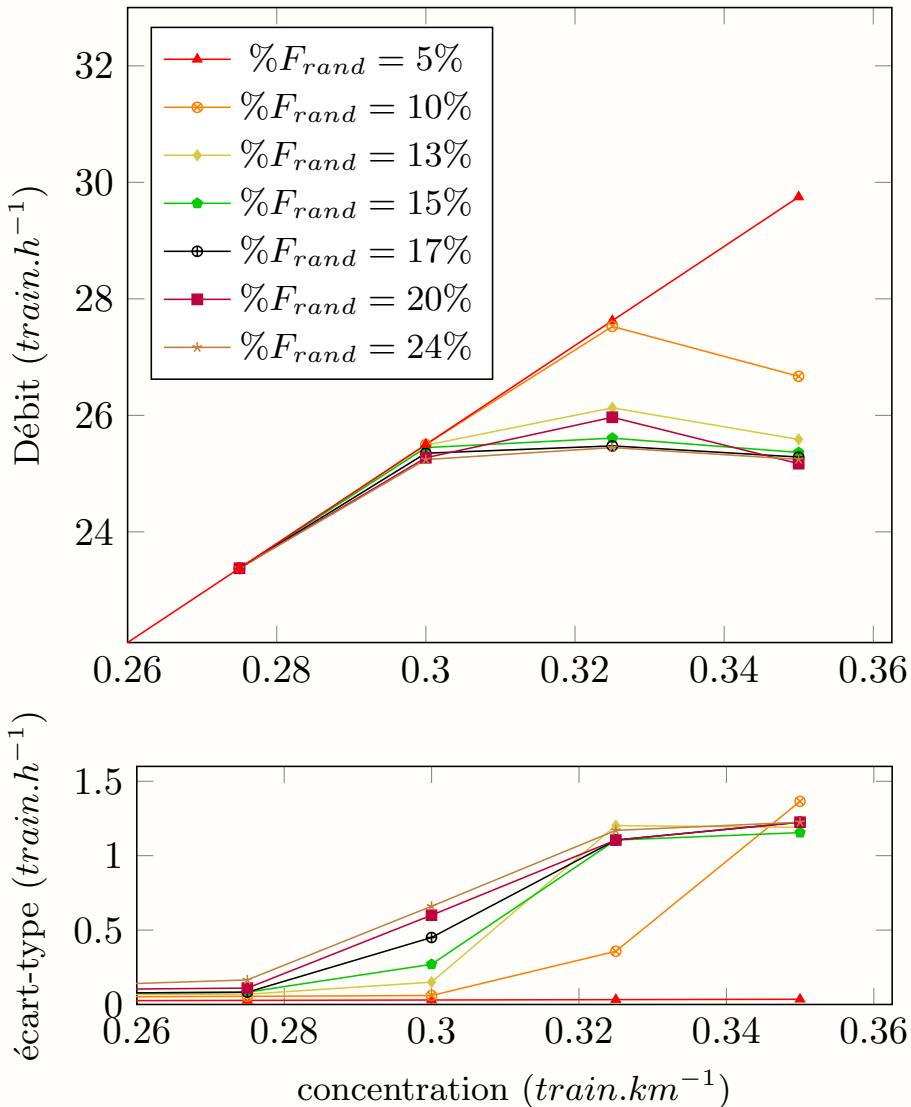
Vitesse moyenne sur les trains puis sur les simulations au cours du temps ainsi que les premiers et derniers quartiles de la distribution des vitesses moyennes sur les trains. Les zones tracées représentent donc 50% des simulations. **a** présente les cas de figure où le couple (T, v_{max}) vaut $(13, 90 \text{ km.h}^{-1})$ et où $\%F_{rand}$ prend les deux valeurs extrêmes de 5% et de 24%. **b** présente les cas de figure où le couple $(T, \%F_{rand})$ vaut $(13, 15\%)$ et où v_{max} prend les deux valeurs extrêmes de 86 km.h^{-1} et de 110 km.h^{-1} . **c** présente les cas de figure où le couple $(v_{max}, \%F_{rand})$ vaut $(90 \text{ km.h}^{-1}, 15\%)$ et où T prend les trois valeurs de 12, de 13 et de 14 trains. On remarque sans surprise que la vitesse moyenne $\bar{v}^{T,500}$ diminue et donc que la congestion s'aggrave lorsque l'on augmente le nombre de trains, la force aléatoire ou la vitesse maximum. Dans le même temps la distribution autour de ces vitesses s'élargit aussi.

conservé. Les petites perturbations que subit ce plan de transport ne modifient pas le débit des trains et tous les trains prévus peuvent circuler. En revanche, lorsque le système se trouve dans l'état congestionné dégradé, le débit devient inférieur au débit planifié et on ne peut faire circuler le trafic désiré.

Figure 2.16 - Diminution de la vitesse moyenne en congestion



Diminution de la vitesse moyenne en congestion. On a repris la figure 2.14 que l'on a enrichie d'une quantification de la congestion. La grille bleue de la sous-figure **a** représente les cas où $v_{\max} = 90 \text{ km.h}^{-1}$. La grille rouge de la sous-figure **b** représente les cas où $a_{r \max}/a_{\max} = 15\%$. Les unités du triplé des paramètres sont ($v_{\max}[\text{km.h}^{-1}]$, $a_{r \max}/a_{\max}[\%]$, T). Pour chaque point de l'espace des paramètres pour lesquels la congestion apparaît, on a calculé $\bar{v}^{T,500,t}$ sur la seconde demi-heure de simulation et on a affiché la baisse de vitesse en km.h^{-1} par rapport à la vitesse planifiée qui est de 85 km.h^{-1} . On remarquera que pour certains points entourés en vert, la congestion fait bien son apparition sans que $\bar{v}^{T,500,t}$ ne soit impactée. Certes le système se congestionne et certains trains prennent du retard mais ces trains peuvent compenser ce retard un peu plus tard en roulant plus vite. C'est l'état du trafic de congestion résiliente.

Figure 2.17 - DFLF


Ce diagramme fondamental de ligne ferroviaire représente le débit ferroviaire en fonction de la concentration pour $v_{max} = 90 \text{ km} \cdot h^{-1}$ et pour une intensité de la force aléatoire et du nombre de trains variant respectivement entre 5% et 24% d'une part et entre 1 et 14 d'autre part. Ce débit ferroviaire est estimé à partir de la vitesse moyenne $\bar{v}^{T,500,t}$, on trace également l'écart type associé (celui de la distribution des $\bar{v}^{T,t}$).

Chapitre 3

Prise en compte des gares

Sommaire

3.1 Évolution du modèle pour prendre en compte les gares	82
3.1.1 L'infrastructure et le plan de transport	82
3.1.2 Implémentation numérique du modèle	88
3.2 Apparition de la congestion	90
3.2.1 Organisation des simulations	90
3.2.2 Résultats	91
3.3 Quantification de la congestion	98
3.3.1 Organisation des simulations	98
3.3.2 Dynamique de la vitesse moyenne instantanée	98
3.3.3 Diagramme Fondamental de Ligne Ferroviaire	101
3.4 Conclusions sur le modèle développé	104
3.4.1 Aspects statiques	104
3.4.2 Aspects dynamiques	105

3.1 Évolution du modèle pour prendre en compte les gares

Une des critiques que l'on peut faire sur le modèle développé dans le chapitre précédent porte sur l'hypothèse d'absence de gares. Les travaux de cette thèse sont financés par Transilien et l'étude présentée dans le chapitre 4 porte d'ailleurs sur une partie de ce réseau. Dans la mesure où c'est un réseau urbain et qu'il comporte donc de nombreuses gares, il convient d'étudier comment le modèle du chapitre précédent peut être modifié par la prise en compte des gares, et donc par là même, par la prise en compte d'une vitesse non homogène.

On a pour cela défini une infrastructure composée d'une succession de cantons dont certains sont des gares. Les tailles et les vitesses de circulation sur ces cantons sont adaptées aux phases de freinage et d'accélération induites par la présence des gares et respectent les distances de freinage minimum exigées par la sécurité.

Après avoir défini une telle infrastructure, on présentera les quelques règles supplémentaires qu'il faut ajouter à notre modèle pour qu'il intègre les phases d'accélérations, de freinage ainsi que les arrêts en gare. Enfin, avant de faire fonctionner le modèle et d'en étudier les résultats, on explicitera les difficultés qu'imposent ces nouvelles règles sur l'implémentation numérique du modèle et on expliquera comment les surmonter. Le modèle développé dans ce chapitre n'est qu'une variante plus complexe du modèle décrit dans le chapitre précédent. Le déplacement des trains repose donc toujours sur les quatre mêmes forces : F_{inter} , $F_{v\ max}$, $F_{horaire}$ et F_{rand} .

3.1.1 L'infrastructure et le plan de transport

Considérons comme en 2.2.1 une ligne de train orientée dans un seul sens et bouclée sur elle-même. Cette ligne est constituée d'une répétition de M motifs. Un motif étant composé lui-même d'une succession de quatre cantons a, b, c et g , auxquels on attribue une signalisation à trois aspects :

- Le canton a vient après une gare, c'est un canton d'accélération.
- Le canton b est un canton où le train roule à vitesse constante.
- Le canton c est un canton où le train freine à l'approche d'une gare.
- Le canton g contient une gare.
- Le motif $abcg$ se répète et le nombre de cantons C vaut : $C = 4 \times M$.

La longueur de ces cantons sera définie un peu plus loin, on discutera en 3.2.2 de la possibilité d'utiliser un motif différents. À partir de ce motif, on doit construire un plan de transport. Ce dernier est moins évident à définir que le plan de transport défini au chapitre précédent. Il faut prendre en compte les accélérations, les freinages et les arrêts en gare. Il faut dans le même temps faire attention à ce que les distances de freinage minimum soient respectées¹.

1. On rappelle qu'il faut simplement qu'un train puisse s'arrêter sur la longueur d'un canton sur lequel il est entré à la vitesse maximale autorisée.

On fait le choix de prévoir les phases d'accélération et de freinage du plan de transport avec une valeur de l'accélération inférieure en valeur absolue au a_{max} . On a : $|a_{plan}| < a_{max}$. Ce choix a la justification suivante :

Pour un train en phase d'accélération et en retard, l'accélération de la force horaire ne peut, par construction, (voir équation (A.4) dans l'annexe A) dépasser la valeur de a_{max} . Ainsi le choix de prendre $|a_{plan}|$ plus petit que a_{max} doit permettre aux trains de rattraper un retard même en phase d'accélération.

Pour un train en phase de freinage et en avance, la justification du choix de a_{plan} est la même. On veut que le train puisse, au moyen de la force horaire, contrecarrer son avance. Cependant le besoin de s'arrêter avec toute la longueur du train à quai est une contrainte supplémentaire qui va brider le comportement du conducteur dans ce dernier cas. On considère en effet, que c'est une phase où le conducteur est concentré pour ne pas rater le quai. Cette concentration particulière du conducteur lors de cette phase peut être modélisée en définissant de nouvelles règles dans le modèle qui de fait vont venir brider la force aléatoire et la force horaire. On choisit par souci de simplicité la même valeur de $|a_{plan}|$ pour les phases de freinage et celles d'accélération. Les nouvelles règles qui encadrent les phases de freinage dans le modèle sont définies de la manière suivante :

- La vitesse de marge est nulle dans la phase de freinage pour imposer à la tête du train de ne pas dépasser le bout du quai.
- Sauf à obéir à un signal rouge, un train ne pourra pas rouler plus lentement que le minimum entre la vitesse de sécurité et la vitesse planifiée. Ceci assurant à la queue du train de bien s'arrêter sur le quai.

Nous n'avons pas exprimé ces règles supplémentaires sous forme de forces. Bien que ce soit possible cela n'est pas d'un grand intérêt : en effet, la complexité croissante de notre modèle, avec l'ajout des gares, rend une description sous forme de forces nettement moins transparente pour le lecteur. De plus l'intérêt d'exprimer le système sous forme de forces est de pouvoir conduire plus simplement des analyses physiques comme c'est le cas dans l'annexe C avec l'analyse de la compétition des effets entre la force horaire et la force aléatoire. Cette analyse reste d'ailleurs valable pour éclairer le comportement du modèle modifié par l'ajout des gares. Les systèmes habituellement étudiés en ferroviaire présentent une complexité de comportement de conduite encore supérieure à celle développée dans ce chapitre. Cela explique pourquoi la modélisation sous forme de forces n'y est pas habituelle. Notre modèle devient donc un modèle mixte avec certaines règles implémentées sous forme de forces, ce qui facilite les études analytiques et d'autres qui ne le sont pas.

La figure 3.1 présente sur la longueur d'un motif la vitesse et l'accélération prévues par le plan de transport mais aussi les règles supplémentaires en termes de marges de vitesses dont on vient de parler. v_{max} permet de voir que la marge de vitesse devient nulle sur les phases de freinage et v_{min} de voir la vitesse minimum imposée sur ces mêmes phases. On remarquera que les limites des cantons ne sont pas encore définies sur cette figure.

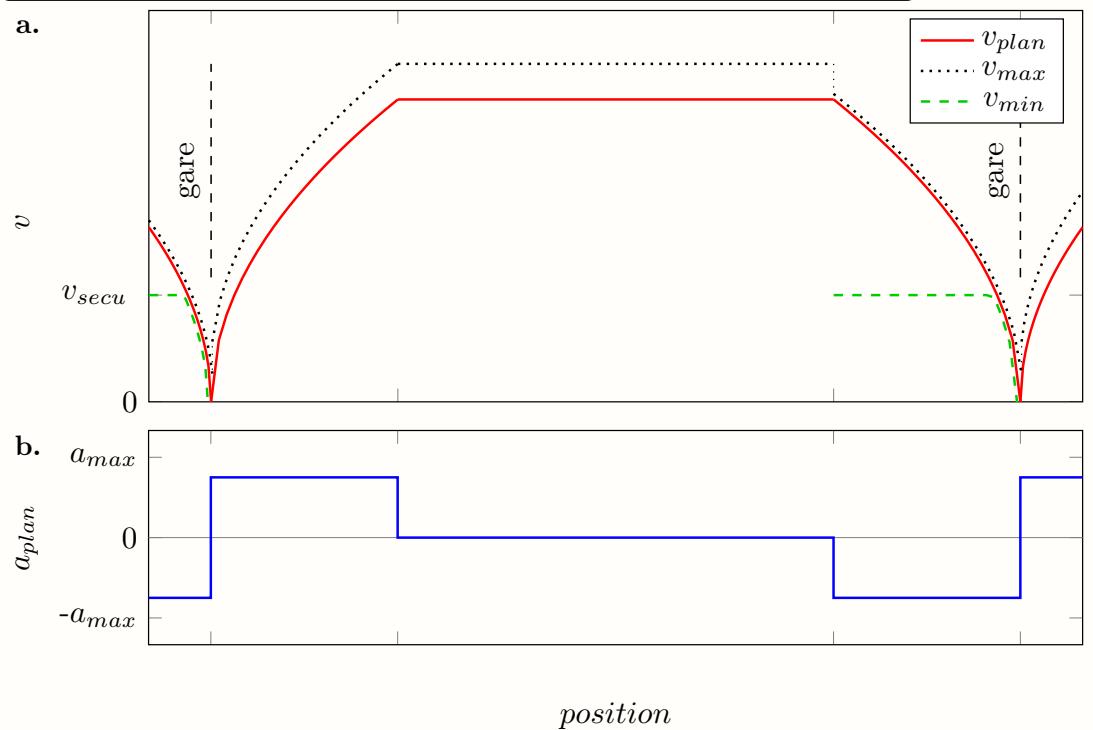
La figure 3.1 ne présente pas à proprement parler le plan de transport mais la marche d'un seul train. Afin de construire le plan de transport dans son intégralité, il faut trouver

la règle d'espacement entre les trains à appliquer. On va voir au moment de dimensionner les cantons que le plan de transport le plus robuste consiste en un espacement des trains homogène ce qui est équivalent au critère de distance entre les trains lorsque la vitesse est constante comme c'est le cas dans le modèle sans gare.

La gestion des arrêts

Le plan de transport prévoit en plus de la marche des trains présentée sur la figure 3.1, un temps de stationnement en gare $t_{stop\ plan}$. Concrètement, lorsqu'un train arrive en gare il déclenche un compte à rebours et ne peut repartir que lorsque ce dernier atteint zéro. Aucune force ne s'exprime durant cet arrêt, ce qui signifie que la force aléatoire ne crée pas d'aléa sur les temps d'arrêt en gare. Pourtant de tels aléas existent bel et bien et peuvent par exemple être dus aux comportements des passagers sur le quai. En pratique, nous al-

Figure 3.1 - Planification de la marche d'un train entre deux gares



Planification de la marche d'un train entre deux gares et limites de vitesse associées. Ces graphiques ne sont pas à l'échelle. Les limitations des cantons ne sont pas représentées. La sous-figure a représente la vitesse planifiée du train entre deux gares ainsi que les limitations de vitesse. La vitesse maximum est obtenue en ajoutant une marge de vitesse constante à la vitesse planifiée, sauf dans la phase de freinage où cette marge est nulle. La vitesse minimum n'est définie que dans la phase de freinage et vaut le minimum entre la vitesse de sécurité et la vitesse planifiée. Ces précautions lors de la phase de freinage assurent que l'arrêt du train coïncide bien avec le quai. Il faut cependant noter qu'en cas d'approche d'un signal rouge, v_{min} ne s'applique plus. La sous-figure b représente les valeurs de l'accélération planifiée entre deux gares.

lons tester et comparer les effets générés par la force aléatoire, telle qu'elle est déjà définie dans le chapitre précédent, à ceux générés par une variation aléatoire du temps d'arrêt en gare t_{stop} autour de $t_{stop plan}$. Nous ne testerons toutefois pas dans ce manuscrit la combinaison des deux sources d'aléas. Cette combinaison n'est pas dénuée d'intérêt, mais nous avons fait le choix d'étudier en priorité les effets des deux types d'aléas pris séparément.

Les règles de définition de t_{stop} sont les suivantes : à chaque arrêt en gare on tire de manière uniforme un temps aléatoire t_r dans l'intervalle $[0, t_{r max}]$ et on pose $t_{stop} = t_{stop plan} + t_r$. Le tirage s'est renouvelé pour chaque train dans chaque gare. L'hypothèse d'uniformité est avant tout un choix de simplicité. Le temps d'arrêt ainsi défini ne peut pas descendre en dessous de $t_{stop plan}$. C'est un choix réaliste, les temps de stationnement prévus par un plan de transport sont rarement ou peu compressibles. Dans ce chapitre on continue de s'intéresser à l'effet de petites perturbations sur le trafic, c'est pourquoi on impose une borne supérieure au temps de stationnement. Mais contrairement à la force aléatoire qui est physiquement bornée, aucune limitation n'interdit de prendre une borne arbitrairement grande pour $t_{r max}$ dans le cadre d'autres travaux. Notons que, contrairement à F_{rand} , t_r n'est pas corrélé avec lui-même au cours du temps.

On remarquera enfin que ce modèle de détermination du temps de stationnement ne prend pas en compte l'accumulation éventuelle de passagers sur le quai et dans les rames, qui ne manquerait pas d'augmenter le temps de stationnement, et qui serait dû à un espace-ment temporel trop grand entre deux trains. La prise en compte de l'accumulation des passagers pourrait d'ailleurs modélisé aussi bien de manière déterministe que de manière aléatoire.

Le dimensionnement des cantons

Dans la réalité, le positionnement des signaux qui délimitent les cantons obéit à quelques contraintes : tout d'abord la position des gares impose la position de certains cantons, ensuite certains éléments d'infrastructure comme les aiguillages imposent également l'implantation de signaux, enfin les distances de freinage doivent être respectées². Un cantonnement qui respecte ces règles n'est pas pour autant un «bon» cantonnement. Parmi les choix de cantonnement possibles, il convient de regarder et de tenter d'optimiser les deux critères suivants :

- Coût : Implanter un signal coûte cher, cela est dû à la fois au fait que des travaux sur une voie de chemin de fer coutent plus cher qu'ailleurs et au fait que les systèmes implantés ont des standards de fiabilité très élevés. Pour faire simple on essaye de ne pas planter plus de signaux que nécessaire.
- Débit : La position des signaux doit être choisie, quand cela est possible, de manière à maximiser la capacité d'une ligne prise dans son ensemble.

Évidemment, en raison des contraintes énoncées ci-avant et dans la mesure où des signaux sont souvent déjà implantés sur la voie, une évolution vers une situation optimale

2. Ce sont les règles de sécurité qui imposent de respecter des distances de freinage réalisables dans les situations les plus défavorables, et ce afin de prévenir de toutes collisions.

n'est pas toujours possible. Nous allons néanmoins dans notre modèle faire l'hypothèse que notre cantonnement maximise la capacité de la ligne tout en minimisant le coût c'est-à-dire en diminuant au maximum le nombre de signaux implantés au kilomètre.

Pour mener à bien cette optimisation, nous utiliserons la formule (1.2). Cette formule donne le débit maximum autorisé par une paire de cantons dans le cas où la vitesse de circulation n'est pas constante. En pratique la quantité déterminante est la somme du temps de visibilité du signal $t_p(d_v)$, du temps de circulation nécessaire pour parcourir deux cantons successifs $t_p(l_{c1}) + t_p(l_{c2})$, et du temps nécessaire pour que le dernier essieu du train quitte le second canton $t_p(l_t)$.

$$D_{max} = 1/t_p = 1/(t_p(d_v) + t_p(l_{c1}) + t_p(l_{c2}) + t_p(l_t))$$

En pratique comme dans le chapitre précédent nous allons négliger le temps de visibilité du signal $t_p(d_v)$. Il est important de le répéter : ce temps, qui a une valeur typique de 35s chez Transilien, est loin d'être négligeable dans le calcul de la capacité d'un système réel. Mais ce qui nous intéresse dans cette thèse, c'est d'étudier comment les aléas de conduite des trains interagissent avec un cantonnement fixe associé à une signalisation à trois aspects, et l'impact de cette interaction sur le trafic. Il se trouve que la prise en compte du temps de visibilité du signal ne change pas qualitativement les effets qu'a le cantonnement sur le trafic et ne fait que diminuer la capacité de la ligne et accélère l'apparition des phénomènes de congestion lorsque l'on augmente le nombre de trains sur cette ligne. Or, dans la mesure où à aucun moment nous ne calibrons les valeurs des paramètres de notre modèle sur un cas d'étude, c'est en premier lieu le caractère qualitatif des résultats observés qui importe et dans une moindre mesure le caractère quantitatif mais de manière grossière. Dans ce contexte nous nous permettons, afin de ne pas complexifier le raisonnement, de négliger le temps de visibilité du signal et de considérer que l'effet de ces signaux s'applique à leur franchissement. Donc l'équation 1.2 devient :

$$D_{max} = 1/t_p = 1/(t_p(l_{c1}) + t_p(l_{c2}) + t_p(l_t)) \quad (3.1)$$

Cette formule donne le débit maximum autorisé par une suite de deux cantons. Pour calculer le débit général d'une ligne, il faut évaluer le débit de toutes les paires de cantons successifs de cette ligne, le débit maximum de la ligne est limité par le plus petit débit maximum obtenu localement. Ainsi, si une paire de cantons présente un débit maximum plus faible que d'autres paires de cantons alors la situation n'est pas optimale. On peut alors, selon les contraintes auxquelles on obéit, soit :

1. Augmenter le débit de cette première paire de cantons afin d'augmenter la capacité de la ligne, ce qui s'obtient en diminuant le temps de parcours sur cette paire de canton. Et comme on roule déjà plus ou moins à la vitesse maximum, il faut diminuer la taille de cette paire de cantons.
2. Diminuer le débit des autres paires de cantons, on ne modifie pas alors la capacité de la ligne mais, en augmentant la longueur et donc le temps de parcours sur ces cantons, on diminue le nombre de cantons par kilomètre et donc le coût de l'installation. Cela ne diminue pas le débit maximum de la ligne car celui-ci est déjà contraint par la paire de canton la plus pénalisante.

Souvent le temps d'arrêt en gare augmente très fortement les temps de parcours des paires de cantons qui contiennent une gare. Ce sont donc ces paires de cantons qui vont limiter le débit de l'ensemble de la ligne. Or, comme le temps d'arrêt en gare limite la diminution du temps de circulation sur ces paires de cantons, nous allons opter pour la seconde solution.

Si on revient à notre infrastructure composée de M répétitions du motif de canton $abcg$ alors l'optimisation du placement des cantons en termes de capacité comme en termes de prix, selon ce que l'on vient de dire et en accord avec l'équation (3.1), revient à résoudre le système d'équations suivant où $t_p(ij+)$ représente le temps de parcours des cantons consécutifs i et j additionné du temps nécessaire pour que le dernier essieu du train quitte le canton j :

$$t_p(ab+) = t_p(bc+) = t_p(cg+) = t_p(ga+) \quad (3.2)$$

On peut écrire ce système sous forme paramétrique en écrivant une équation et une variable supplémentaire :

$$(3.2) \Leftrightarrow \begin{cases} t_p(ab+) = t \\ t_p(bc+) = t \\ t_p(cg+) = t \\ t_p(ga+) = t \end{cases}$$

On va décomposer dans un premier temps chacun de ces termes en faisant les hypothèses suivantes sur les cantons :

- Le canton a comporte une section³ d'accélération a_1 puis une section parcourue à vitesse constante a_2 .
- Le canton b comporte une section à vitesse constante uniquement b_1 .
- Le canton c comporte une section à vitesse constante c_1 puis une section de décélération c_2 .
- Le canton g comporte une section de décélération g_1 , une section d'arrêt g_{stop} de taille nulle mais de temps d'arrêt t_{stop} non nul. Et une section d'accélération g_2 .

On décompose ensuite les premières sections (indice 1) de chaque canton $i \in a, b, c, g$ en une première sous-section dont la longueur est égale à la taille d'un train $i_{1,1}$ et une seconde sous-section $i_{1,2}$ qui complète la section i_1 . La première sous-section va permettre d'exprimer le temps qu'il faut pour que le dernier essieu d'un train quitte le canton précédent. La décomposition en segments de canton qui en résulte est visible dans le tableau 3.1. Le système devient :

3. On parle ici des sections prévues par le plan de transport.

$$(3.2) \Leftrightarrow \begin{cases} t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + t_p(b_{1,1}) + t_p(b_{1,2}) + t_p(c_{1,1}) \\ \quad t_p(b_{1,1}) + t_p(b_{1,2}) + t_p(c_{1,1}) + t_p(c_{1,2}) + t_p(c_2) + t_p(g_{1,1}) \\ \quad t_p(a_{1,1}) \\ \quad + t_p(c_{1,1}) + t_p(c_{1,2}) + t_p(c_2) + t_p(g_{1,1}) + t_p(g_{1,2}) + t_{stop} + t_p(g_2) = t \\ t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + t_p(b_{1,1}) \\ \quad + t_p(g_{1,1}) + t_p(g_{1,2}) + t_{stop} + t_p(g_2) = t \end{cases} = t$$

Une fois les cantons dimensionnés de manière à vérifier cette équation les paires de trains ne peuvent pas être espacées temporellement de moins que la valeur de t sans générer une situation de congestion. Si le plan de transport n'espace pas les trains entre eux de manière homogène en termes de temps de parcours, alors nécessairement, certaines paires de trains se retrouvent plus rapprochées (toujours temporellement) que dans la situation homogène. Lors du déroulement du temps de transport, il suffira de perturbations plus petites sur ces paires de trains pour engendrer une situation de congestion par rapport à la situation homogène. Le plan de transport le plus robuste est donc celui où les temps de parcours prévus entre chaque paire de trains consécutifs sont tous égaux.

Application numérique - Notations

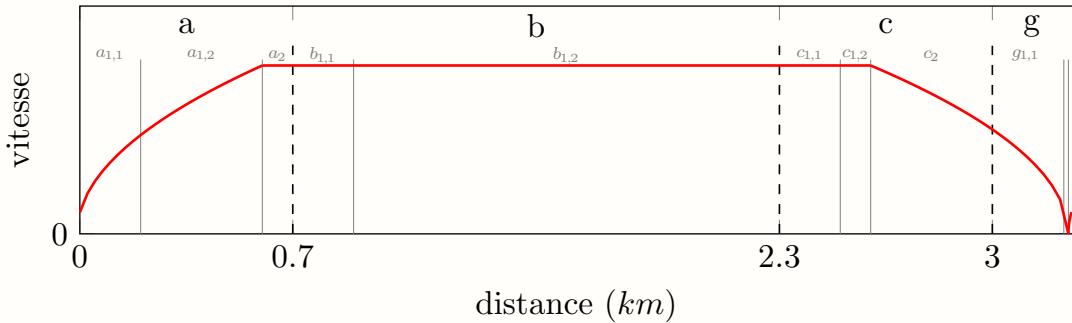
Le système (3.2) est résolu dans l'annexe D. À l'issue de cette résolution analytique, on est capable d'exprimer toutes les longueurs de segments de cantons et tous les temps de parcours correspondants. On fixe les valeurs de certains paramètres et on calcule les autres. La figure 3.2 montre la position des cantons et de tous les segments de cantons par rapport à la marche du train.

On regroupe les tailles et le temps de parcours de chaque segment de cantons dans le tableau 3.1. Le tableau 3.1 présente quant à lui toutes les nouvelles variables introduites dans ce chapitre ou celles dont la valeur utile pour l'application numérique évolue par rapport à sa valeur utilisée au chapitre 2. Pour les variables non-mentionnées il faut se référer au tableau 2.2. Par ailleurs la valeur prise par tous les $t(i, j+)$ vaut 205s. Cela signifie que tant que tous les trains sont temporellement espacés de plus de 205s, la congestion est impossible.

3.1.2 Implémentation numérique du modèle

L'implémentation des nouvelles règles de comportement de conduite et la prise en compte d'une infrastructure plus complexe car non homogène ne sont pas sans poser quelques problèmes lors de la discréétisation du modèle :

- La comparaison pour le calcul de la force horaire entre l'horaire réel et l'horaire planifié pose problème. L'écart horaire n'est en effet, plus proportionnel à une distance. Il faut à chaque fois que l'on veut calculer l'écart à l'horaire d'un train évaluer l'heure qu'il devrait être selon le plan de transport pour que le train soit à la position où il est effectivement. On utilise les formules qui font le lien entre la position sur un segment de canton et le temps passé sur ce même segment. Ces formules ont été démontré dans l'annexe D.

Figure 3.2 - Position des cantons et marche du train


Ce graphique à l'échelle représente la position des cantons et des segments de cantons d'un motif $abcg$ superposés avec la marche du train. Les segments $g_{1,2}$ et g_2 sont trop petits pour être représentés.

- Lors du déplacement du train, les règles auxquelles il est soumis dépendent du canton sur lequel il se trouve mais aussi du canton sur lequel se trouve le dernier essieu du train qui se trouve devant lui. Dans la version précédente du modèle, on pouvait déterminer le numéro de ces deux cantons par une simple division euclidienne de la position absolue par la taille d'un canton. Dans cette nouvelle version du modèle, il faut comparer la distance absolue avec chacune des longueurs présentées dans le tableau 3.1. En revanche à l'échelle du motif, la ligne reste périodique et l'on peut utiliser la division euclidienne ce qui simplifie et accélère le code.
- Les erreurs d'une discréétisation à pas de temps fixe se font sentir sur les arrêts en gares. Dans notre modèle, on «demande» au train de s'arrêter sur une position ponctuelle ce qui n'est pas un problème. Dans la version discréétisée du modèle, il est rare que le train s'arrête exactement sur cette position. Sauf que la gestion de l'arrêt en

Tableau 3.1 - Descriptpion du motif de base de la simulation

canton	a			b			c
segment	$a_{1,1}$	$a_{1,2}$	a_2	$b_{1,1}$	$b_{1,2}$	$c_{1,1}$...
v	/	/	→	→	→	→	...
l (m)	200,0	409,4	118,1	200,0	1387,1	200,0	...
t_p (s)	23,9	21,9	5,0	8,5	58,8	8,5	...
canton	c			g			
segment	...	$c_{1,2}$	c_2	$g_{1,1}$	$g_{1,2}$	<i>stop</i>	$g_{1,2}$
v	...	→	↘	↘	↘	0	/
l (m)	...	118,1	409,4	200,0	10,0	0,0	10,0
t_p (s)	...	5,0	21,9	23,9	6,7	30,0	6,7

Tableau 3.2 - Notations

Nom	Valeur	Description
$a_{plan} (m.s^{-2})$	0,45	accélération planifiée
$v_{plan\ max} (km.h^{-1})$	85	vitesse maximale planifiée
$v_{marge} (km.h^{-1})$	0-5	Marge positive de vitesse autorisée
$l_T (m)$	200	taille d'un train
$t_{stop} (s)$	30	temps d'arrêt
M	15	nombre de motifs
C	60	nombre de cantons
T	15-24	nombre de trains
$l_{abcg} (m)$	3262	longueur d'un motif
$l_{ligne} (km)$	48,9	longueur de la ligne
$t_{r\ max} (s)$	10-30	max du temps d'arrêt additionnel

gare n'est prévue dans le modèle que si le train s'arrête effectivement sur cette position. Il faut donc mettre en place une plage de rattrapage, dont la longueur dépend du pas de temps, capable de gérer les arrêts en gare. La mise en place de cette plage de rattrapage doit être faite avec des précautions particulières car c'est une zone de faible vitesse pour laquelle de petites erreurs de distances peuvent engendrer des erreurs temporelles assez grandes. Une mesure est prévue dans le code pour éliminer ces erreurs en adaptant légèrement le temps d'arrêt en gare.

À l'exception de ces modifications, les éléments présentés dans le chapitre précédent en 2.2.4 restent valables. Le temps de calcul est cependant rallongé par les modifications ci-dessus et en particulier par le remplacement de certaines divisions euclidiennes au profit de tests conditionnels.

3.2 Apparition de la congestion

On garde la même définition de la congestion que précédemment (voir partie 2.3). Notre système est congestionné si F_{inter} est non nulle pour au moins un train. Cette définition de la congestion implique pour notre infrastructure qu'au moins un train doive ralentir à la vitesse de sécurité sur au moins un canton.

3.2.1 Organisation des simulations

Le plan de transport le plus robuste est choisi pour toutes les simulations. C'est le plan de transport pour lequel tous les trains sont espacés du même temps de parcours. On a

montré en 3.1.1 que ce plan de transport était le plus robuste. À l'instant initial de la simulation, tous les trains sont à l'heure et roulent à la vitesse prévue par le plan de transport.

Au chapitre précédent Pour étudier l'apparition de la congestion on faisait varier trois paramètres : la vitesse maximum autorisée v_{max} , l'amplitude des variations de la force aléatoire qui s'écrit $m \cdot a_{r\ max}$ et le nombre de trains T. Le chapitre précédent a montré que l'influence de la vitesse maximum autorisée était moindre par rapport aux deux autres paramètres. On va donc se contenter ici d'étudier l'influence du nombre de trains et des aléas représentés soit par la force aléatoire, soit par des variations du temps d'arrêt en gare.

Variations de la force aléatoire et du nombre de trains

Considérons le couple de paramètres ($a_{r\ max}/a_{max}[\%]$, T) comme les coordonnées d'un point de l'espace des paramètres explorés. Le nombre de trains T dans ces simulations prend toutes les valeurs de 15 à 24 inclus. L'amplitude de la force aléatoire ($a_{r\ max}/a_{max}[\%]$) prend les même valeurs que dans le chapitre précédent à savoir 5%, 10%, 13%, 15%, 17%, 20% et 24%.

Variations du temps de stationnement

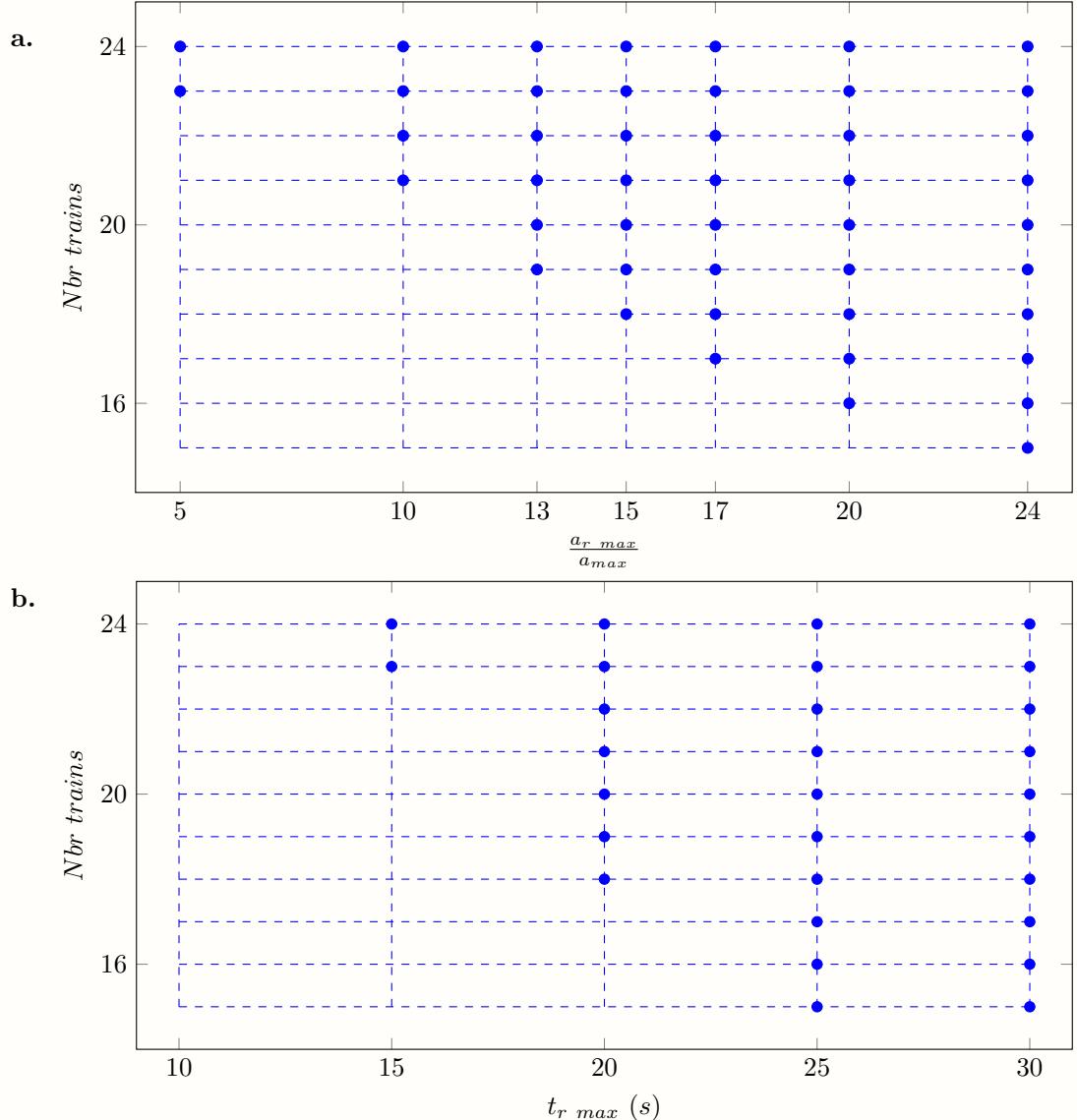
Considérons le couple de paramètres ($t_{r\ max}[s]$, T) comme les coordonnées d'un point de l'espace des paramètres explorés. Le nombre de trains T dans ces simulations prend toutes les valeurs de 15 à 24 inclus. La borne supérieure que peut prendre de manière aléatoire le temps de stationnement additionnel $t_{r\ max}$ peut prendre les valeurs suivantes : 10s, 15s, 20s, 25s et 30s.

Enfin, on réalise les simulations avec M = 15 Motifs de $l_{abcg} = 3262m$, soit C = 60 can-
tons. La longueur des trains est toujours de $l_t = 200m$. L'ensemble des valeurs utiles pour la simulation est résumé dans les tableaux 3.1, 3.2 et 2.2.

3.2.2 Résultats

Pour chaque point de l'espace des paramètres, on mesure la fréquence d'apparition de la congestion au cours du temps. La figure 3.3 montre tous les points de l'espace des paramètres parcouru et distingue ceux pour lesquels on observe l'apparition de la congestion dans un délai de trois heures. Cette figure ne permet pas de distinguer la rapidité avec laquelle la congestion apparaît.

La figure 3.4 représente quelques exemples représentatifs de graphiques d'apparition de la congestion. Les sous-figures **a**, **b** et **c** permettent d'observer l'influence de la variation de la force aléatoire sachant que le temps de stationnement fixe, elles correspondent respectivement à $\%F_{rand} = 5\%$, $\%F_{rand} = 15\%$ et $\%F_{rand} = 24\%$. Les sous-figures **d**, **e** et **f** permettent d'observer l'influence de la variation aléatoire du temps de stationnement sachant l'amplitude de la force aléatoire est nulle, elles correspondent respectivement à $t_{r\ max} = 10s$, $t_{r\ max} = 20$ et $t_{r\ max} = 30s$. Comme attendu, on observe que la congestion augmente avec le nombre de trains et l'amplitude du hasard. Notons que l'amplitude du

Figure 3.3 - Espace des paramètres explorés

Espace des paramètres explorés représenté par le quadrillage. Les cercles représentent les points de l'espace des paramètres pour lesquels la congestion est apparue au moins une fois sur l'ensemble des mille simulations. Sur la sous-figure **a** on représente l'espace des paramètres avec variations de la force aléatoire et sur la sous-figure **b** on représente l'espace des paramètres avec variations aléatoires du temps de stationnement.

hasard a à peu près les mêmes effets que ce hasard s'applique à travers la force aléatoire ou à travers le temps de stationnement. Comme pour le modèle sans gare, les probabilités d'apparition de la congestion restent nulles dans les premiers instants.

Interpolation par notre modèle DTP

L'évolution du modèle présenté dans ce chapitre est toujours un processus de Markov. On va tenter d'interpoler les courbes de probabilité d'apparition de la congestion avec la formulation DTP développée en 2.4.2. On rappelle en quoi consiste cette formulation, on a la probabilité d'apparition de la congestion qui s'écrit :

$$P(t) = 1 - F.e^{S \times t} \cdot \mathbf{1}$$

Avec S une matrice qui prend une forme particulière :

$$S = \begin{pmatrix} -s_0 & s_0 & & 0 \\ & -s_1 & s_1 & \\ & & \ddots & \ddots \\ 0 & & & -s_{(k-1)} \end{pmatrix} ; \quad s_i = \max(0, s_0 + i \cdot \Delta s) \quad \text{avec} \quad \Delta s < 0$$

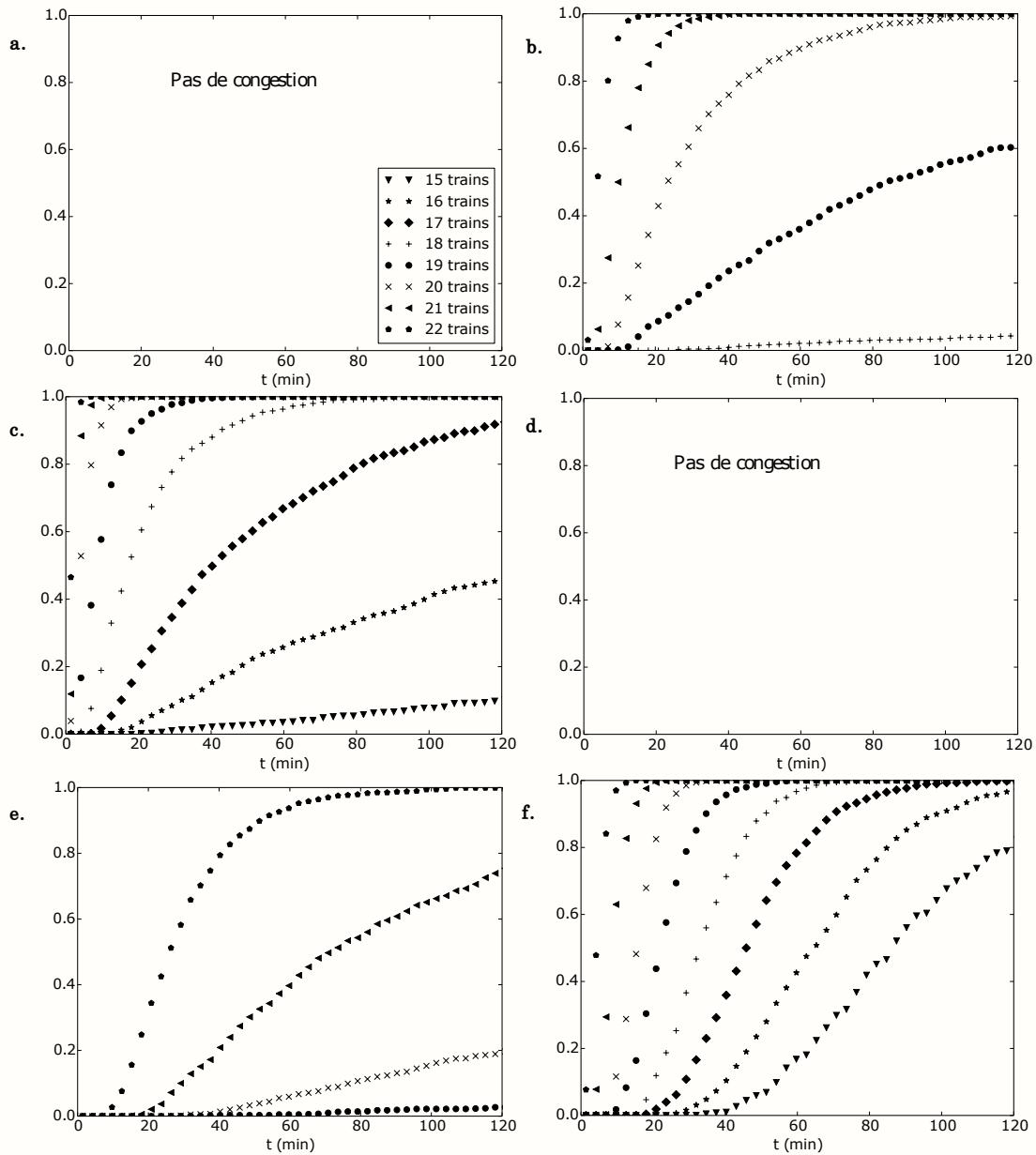
La figure 3.5 présente deux exemples d'interpolation DTP avec $k = 10$ comme taille de la matrice S . L'un de ces exemples est obtenu en utilisant les aléas via la force aléatoire, on a alors choisi de représenter l'exemple où $T = 20$ et $a_r \max / a_{\max} = 15\%$. L'autre exemple est obtenu en utilisant les aléas via les temps de stationnement, on a alors retenu l'exemple où $T = 17$ et $t_r \max = 30s$. Pour toutes les simulations, les interpolations réalisées sont très satisfaisantes.

Rôle de la taille et de l'homogénéité de la ligne

Que ce soit sur la figure 3.4 ou sur la figure 3.5, on remarque que les probabilités prennent une allure générale différente selon que le hasard est implémenté dans le modèle à l'aide de la force aléatoire ou à l'aide des temps d'arrêts. On observe en effet que dans le second cas, les probabilités restent nulles ou proches de zéro plus longtemps que dans le premier. Il y a une raison à cela.

Dans le premier cas les aléas s'expriment sur tous les cantons⁴. En revanche, dans le second cas, non seulement les aléas ne s'expriment qu'en gare, mais en plus de cela ces aléas ne sont pas symétriques.

4. Certes la force s'exprime moins fortement sur les cantons c et g car la phase de freinage obéit à des contraintes supplémentaires.

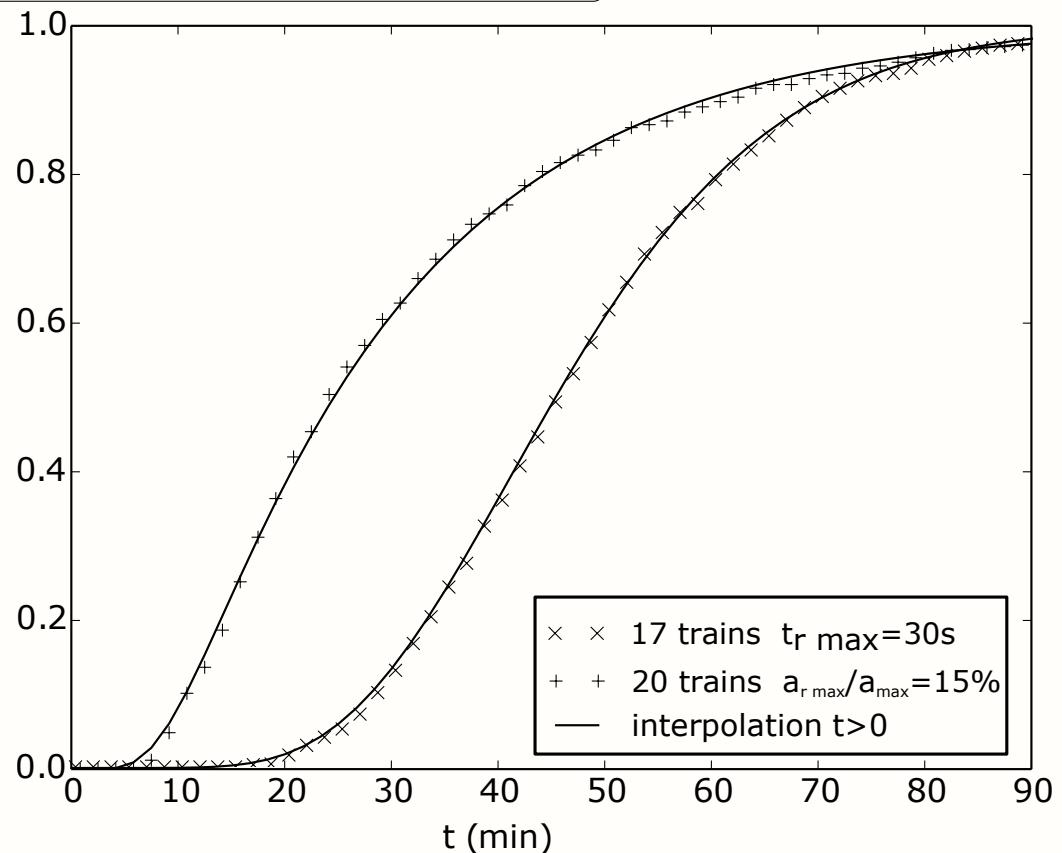
Figure 3.4 - Probabilité d'apparition de la congestion au cours du temps

Probabilité d'apparition de la congestion au cours du temps pour des systèmes comportant 15 à 22 trains et pour 6 couples différents de l'espace des paramètres. Les sous-figures **a**, **b** et **c** permettent d'observer l'influence de la variation de la force aléatoire, elles correspondent respectivement à $\%F_{rand} = 5\%$, $\%F_{rand} = 15\%$ et $\%F_{rand} = 24\%$. Les sous-figures **d**, **e** et **f** permettent d'observer l'influence de la variation aléatoire du temps de stationnement, elles correspondent respectivement à $t_{r max} = 10s$, $t_{r max} = 20$ et $t_{r max} = 30s$. Dans tous les cas la marge de vitesse est de $5km.h^{-1}$. Notons que l'apparition de la congestion ne se fait pas immédiatement et que la probabilité reste nulle dans les premiers instants. Les sous-figures **a** et **d** ne présentent aucune courbe car la probabilité d'apparition de la congestion y est toujours nulle. Dans ces cas, l'amplitude du hasard ne perturbe pas assez le système pour pouvoir générer de la congestion et ce, quel que soit le nombre de trains.

Par exemple, pour un temps d'arrêt additionnel pouvant varier entre 0s et 30s on augmente en moyenne de 15s le temps de stationnement. Cela veut dire que l'on casse l'homogénéité de temps de parcours que l'on avait construite en construisant l'infrastructure, le temps nécessaire pour parcourir deux cantons et sortir du deuxième était indépendant de la paire de cantons considérée : $t(i\ j+) = 126.5s$. Sur les 4 paires de cantons possibles que l'on peut construire à partir du motif $abcg$: ab , bc , cg et ga , les deux paires comprenant le canton g voient leur temps de parcours rallongé de 15s en moyenne. La gare redevient donc, comme c'est souvent le cas dans les systèmes réels, un lieu qui limite le trafic en comparaison avec les autres cantons.

On peut à l'aide d'un raisonnement analytique simple comprendre le rôle que joue la non-homogénéité des temps de parcours sur la probabilité d'apparition de la congestion.

Figure 3.5 - Exemples d'interpolations DTP



Exemples d'interpolations DTP avec $k = 10$ dans notre formulation de la DPT. On rappelle que la vitesse de marge est de $5km.h^{-1}$. On distingue deux exemples d'interpolation sur ce graphique l'un est obtenu en utilisant les aléas via la force aléatoire, on a alors choisi de représenter l'exemple où $T = 20$ et $a_{r\ max}/a_{max} = 15\%$. L'autre exemple est obtenu en utilisant les aléas via les temps de stationnement, on a alors retenu l'exemple où $T = 17$ et $t_{r\ max} = 30s$.

Ce même raisonnement va nous permettre de regarder l'influence d'un autre paramètre que nous n'avons pas étudié dans nos simulations, la taille de la ligne.

Soit une ligne qui comporte C cantons et dont la probabilité de tomber en congestion au cours du temps vaut $P_C(t)$. Si on double la taille de la ligne à l'identique, elle comporte alors $2C$ cantons. Que se passe-t-il pour la probabilité d'apparition de la congestion ?

On fait l'hypothèse que les événements, «la congestion apparaît sur la première moitié de la ligne» et «la congestion apparaît sur la seconde moitié de la ligne» sont indépendants. Cette hypothèse est justifiée si C est grand. Il faut en effet pouvoir négliger le nombre de cantons en contact entre les deux moitiés de ligne sur le nombre de cantons total. On a alors⁵ :

$$P_{2C}(t) = 1 - (1 - P_C(t))^2 = P_C(t) \cdot (2 - P_C(t)) > P_C(t)$$

L'augmentation de la taille de la ligne toutes choses égales par ailleurs implique une augmentation de la probabilité d'apparition de la congestion. En d'autres termes, pour une ligne doublée en longueur, la courbe d'apparition de la congestion sera toujours supérieure à celle issue de la ligne non doublée.

On va maintenant comparer une ligne pour laquelle la congestion a la même probabilité d'arriver en tout point à une ligne où la congestion se focalise en certains points dits limitants.

Dans le premier cas, on fait l'hypothèse que les contributions de chaque paire de cantons sont indépendantes les unes des autres. Encore une fois cette hypothèse est justifiée uniquement si le nombre de cantons est grand. On nomme $p_1(t)$ la probabilité d'apparition de la congestion sur une paire de cantons et $P(t)$ la probabilité d'apparition de la congestion sur l'ensemble de la ligne de C cantons. On a :

$$P(t) = 1 - (1 - p_1(t))^C \quad (3.3)$$

Cette équation est d'ailleurs cohérente avec l'équation précédente. Considérons maintenant une ligne où il y a des points limitants sur laquelle la probabilité de tomber en congestion s'appelle $P_l(t)$. Pour simplifier on considère que C_1 paires de cantons ont une probabilité de voir la congestion apparaître en leur sein $p_1(t)$ et que $C_2 = C - C_1$ paires de cantons ont une probabilité de voir la congestion apparaître en leur sein $p_2(t)$ on a les C_1 paires de cantons qui sont les points limitants donc $p_1(t) > p_2(t)$. On a alors :

$$P_l(t) = 1 - (1 - p_1(t))^{C_1} \cdot (1 - p_2(t))^{C_2} \quad (3.4)$$

Regardons maintenant ce qui se passe au début de nos simulations quand la probabilité totale d'apparition de la congestion est encore faible. La marge de stabilité est plus

5. L'événement : «la congestion apparaît sur la ligne doublé» est la négation de l'événement : «la congestion n'apparaît pas sur la première moitié» ET «la congestion n'apparaît pas sur la seconde moitié». La négation se traduit par $1 - p()$ et le ET se traduit par un produit de probabilité sous réserve qu'elles soient non liés entre elles. Cette écriture est reprise dans les formules suivantes.

faible sur les points limitants elle est donc plus rapidement «grignotés» et $p_1(t)$ devient non nul avant $p_2(t)$. Dans cette situation on a :

$$p_1(t) \rightarrow 0^+ \Rightarrow p_2(t) = 0$$

donc : $P_l(t) = 1 - (1 - p_1(t))^{C_1}$

Dans cette situation on pose $p_1(t) = x \rightarrow 0^+$ et on calcule le rapport $P_l(t)/P(t)$ qui va nous permettre de savoir si une des deux fonctions écrase l'autre au voisinage de ce point :

$$\frac{P_l(t)}{P(t)} = \frac{1 - (1 - x^{C_1})}{1 - (1 - x^C)}$$

$$\frac{P_l(t)}{P(t)} = \frac{C_1 \cdot x + \dots}{C \cdot x + \dots} \sim \frac{C_1}{C} < 1$$

$P_l(t)$ s'éloigne donc C/C_1 fois moins vite de zéro que $P(t)$. On en conclut que la dynamique d'apparition de la congestion est affectée par l'homogénéité de la ligne. On observe en effet sur la figure 3.4 où sur la figure 3.5, dans le cas où l'aléa est modélisé par les temps d'arrêts, c'est-à-dire le cas avec des points limitants, que les probabilités d'apparition de la congestion augmentent plus lentement au cours du temps.

Pour étudier le rôle du manque d'homogénéité de la ligne et se rapprocher d'un cas réel, on peut par exemple reprendre le modèle avec gares en effectuant quelques modifications dans l'infrastructure. En remplaçant quelques motifs *abcg* par des motifs *abbcg* par exemple. Le programme nécessiterait une centaine de lignes de codes supplémentaires pour prendre en compte une telle modification, car en brisant la périodicité des motifs, on perd certaines simplifications dans le code. Ce travail n'a pas été réalisé mais on peut néanmoins à l'aide du raisonnement précédent décrire qualitativement comment la congestion se formerait sur une telle infrastructure.

Un motif *abbcg* introduit de nouvelles combinaisons de paires de cantons. On perd l'homogénéité des temps de parcours $t(ij+)$ qui faisait tous 126,5s. Désormais on a un temps plus long que les autres : $t(bb+) = 142,9s$. La première conséquence de cette hétérogénéité est l'obligation de concevoir un plan de transport en respectant une nouvelle contrainte : le temps entre deux trains doit être supérieur à 142,9s au lieu de 126,5s précédemment. Le débit maximum de la ligne diminue donc. Si le temps entre deux trains est par exemple de 200s et que les aléas sont capables, dans la pire des situations, de rapprocher deux trains de 50s, alors le temps entre deux trains ne pourra jamais être inférieur à 150s et la congestion ne pourra pas apparaître.

En revanche, si l'on augmente soit l'amplitude des aléas, soit le nombre de trains, on va rencontrer des situations pour lesquelles le temps qui sépare deux trains va être inférieur à 142,9s. Seulement, il faut que cet événement se produise spécifiquement au moment où ces deux trains se trouvent sur le motif *abbcg*, la probabilité que cela se produise est donc moins élevée. C'est pour cela que la dynamique de la congestion est plus lente dans les situations hétérogènes. Mais comme on vient de le dire, dans la même situation, la congestion serait impossible sur l'infrastructure homogène avec des motifs *abcg*, et il faudrait encore augmenter le nombre de trains pour que cela se produise.

3.3 Quantification de la congestion

3.3.1 Organisation des simulations

On va maintenant quantifier la congestion dont on vient d'étudier l'apparition. Les seuls changements dans l'organisation des simulations sont mineurs : le temps des simulations passe à une heure et la simulation ne prend plus fin lorsque la congestion apparaît. Toutes les vingt secondes de temps simulé, le programme calcule la vitesse moyenne instantanée sur l'ensemble des trains du système. On note cette vitesse \bar{v}^T . Le nombre de simulations dans la méthode de Monte-Carlo a été ramené à seulement cinq cents.

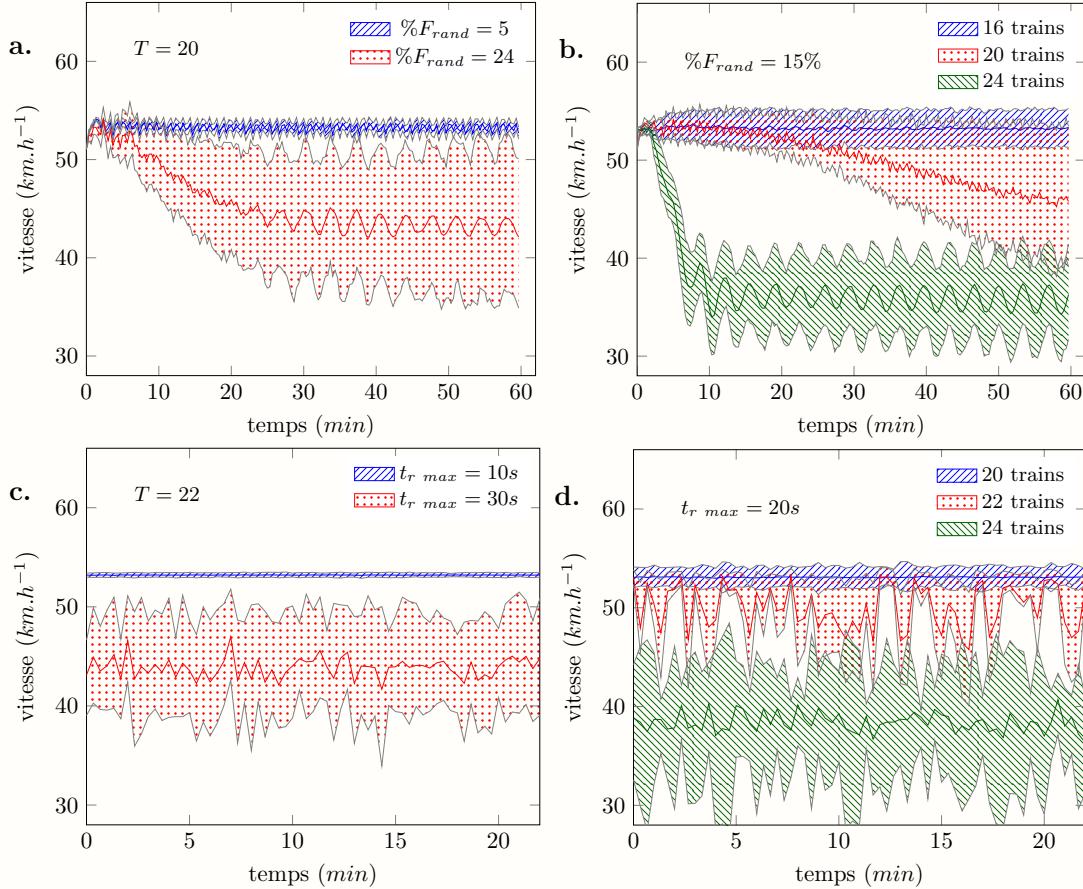
3.3.2 Dynamique de la vitesse moyenne instantanée

À partir des vitesses moyennes instantanées de chaque simulation, on calcule à nouveau une moyenne sur l'ensemble des cinq-cents simulations. On note cette vitesse $\bar{v}^{T,500}$. La différence entre cette vitesse et la vitesse planifiée est donc un indicateur de quantification de la congestion. On calcule également le premier et le dernier quartile de la distribution de \bar{v}^T en fonction du numéro de la simulation. Quelques graphiques les plus significatifs présentent ces résultats dans la figure 3.6. Les sous-figures **a** et **b** représentent des exemples dans lesquels le hasard est implémenté à l'aide de la force aléatoire alors que les sous-figures **c** et **d** représentent des exemples dans lesquels le hasard est implémenté à l'aide du temps de stationnement. L'absence de régime transitoire dans ces deux derniers cas est due au fait que l'aléa s'applique dès le début de la simulation alors que dans le cas de la force aléatoire, il faut laisser du temps pour que les forces agissent.

On voit de petites vagues périodiques apparaître sur les sous-figures **a** et **b**, c'est un phénomène normal qui est dû à une certaine synchronisation des arrêts en gare. Avec une ligne de 15 motifs de long, si on avait représenté la vitesse moyenne pour 15 trains, elle serait périodiquement nulle puisque les arrêts seraient parfaitement synchronisés. L'analyse que l'on a présenté en 2.5.2 pour expliquer le passage du régime transitoire au régime permanent reste valable pour les sous-figures **a** et **b**. Les explications sur le rôle de l'amplitude des aléas et du nombre de trains sur la vitesse moyenne restent valables pour toutes les sous-figures.

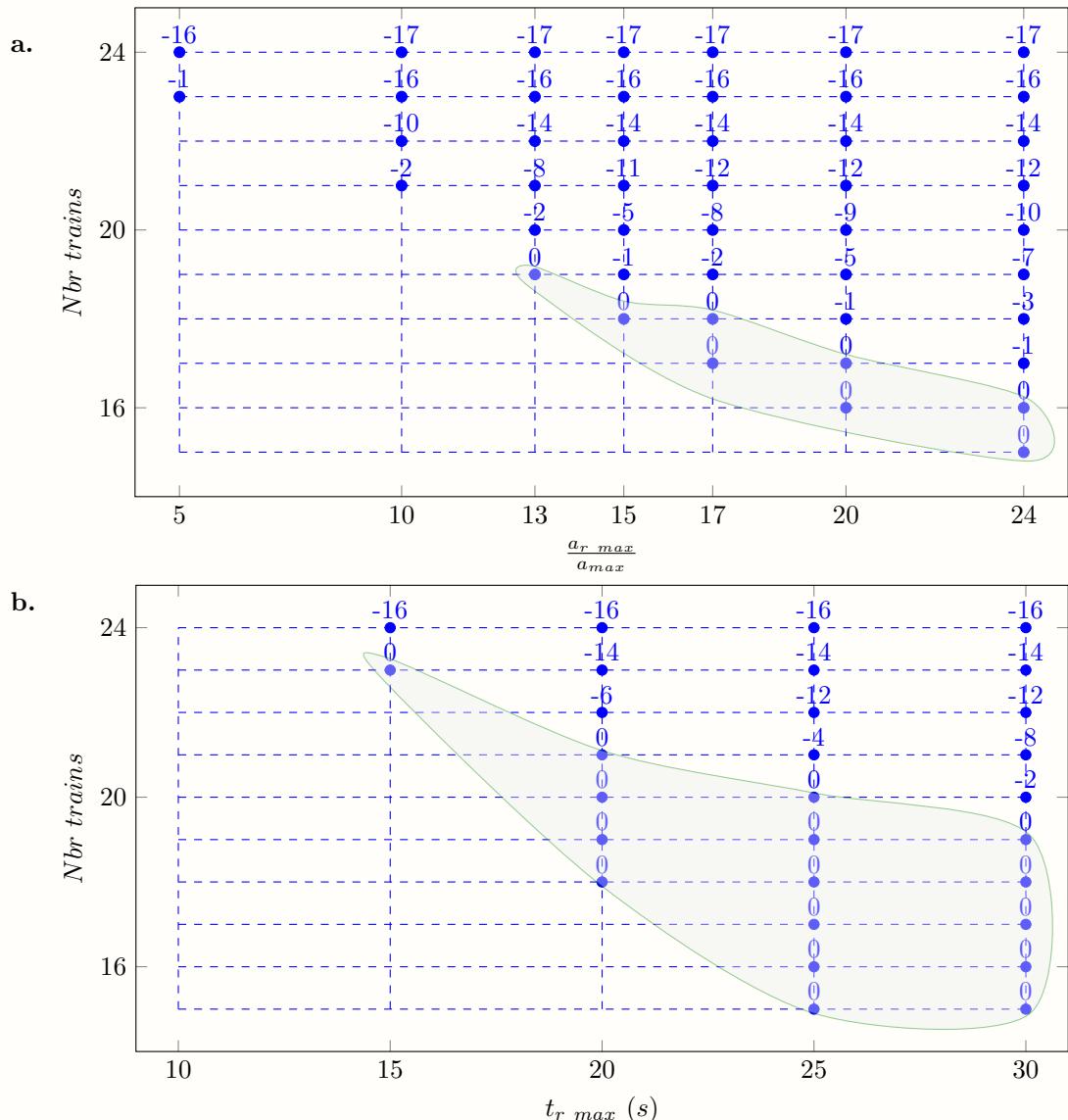
On remarque comme au chapitre précédent que la congestion est parfois de faible gravité au point que la vitesse moyenne des trains $\bar{v}^{T,500}$ égale la vitesse planifiée. Dans ces cas-là l'indicateur que l'on a construit en définissant différemment l'objectif de performance représente une amélioration notable mais n'est toujours pas parfait : il est évident que de petits retards locaux sont observés et que la qualité du service est légèrement dégradée même si ces retards sont ratrappés par la suite. On remarque sans surprise que la vitesse moyenne $\bar{v}^{T,500}$ diminue et donc que la congestion s'aggrave lorsque l'on augmente le nombre de trains ou l'amplitude des aléas. Dans le même temps la distribution autour de ces vitesses s'élargit aussi.

La figure 3.7 reprend la figure 3.3 en ajoutant une information sur la quantification de la congestion : on représente pour chaque point de l'espace des paramètres pour lequel

Figure 3.6 - $\bar{v}^{T,500}(t)$


Vitesse moyenne sur les trains puis sur les simulations au cours du temps ainsi que les premiers et derniers quartiles de la distribution des vitesses moyennes sur les trains. Les zones tracées représentent donc 50% des simulations. **a** présente les cas de figure où le couple (T, t_{r_max}) vaut $(13,0s)$ et où $\%F_{rand}$ prend les deux valeurs extrêmes de 5% et de 24%. Le cas où $\%F_{rand} = 15\%$ peut être retrouvé dans la sous-figure **b**. **b** présente les cas de figure où le couple $(t_{r_max}, \%F_{rand})$ vaut $(0s, 15\%)$ et où T prend les trois valeurs de 16, de 20 et de 24 trains. **c** présente les cas de figure où le couple $(T \%F_{rand})$ vaut $(13,0\%)$ et où t_{r_max} prend les deux valeurs extrêmes de 10s et de 30s. Le cas où $t_{r_max} = 20s$ peut être retrouvé dans la sous-figure **d**. **d** présente les cas de figure où le couple $(t_{r_max}, \%F_{rand})$ vaut $(20s, 0\%)$ et où T prend les trois valeurs de 20, de 22 et de 24 trains. L'échelle de temps n'est pas la même dans ces deux derniers graphiques pour permettre de visualiser des variations plus rapides dans le cas d'un aléa sur les temps de stationnement.

Figure 3.7 - Diminution de la vitesse moyenne en congestion



Diminution de la vitesse moyenne en congestion. On a repris la figure 3.3 que l'on a enrichie d'une quantification de la congestion. La sous-figure **a** représente les cas où $t_{r\ max} = 0s$. La sous-figure **b** représente les cas où $a_{r\ max}/a_{max} = 0\%$. Pour chaque point de l'espace des paramètres pour lesquels la congestion apparaît, on a calculé $\bar{v}^{T,500,t}$ sur la seconde demi-heure de simulation et on a affiché la baisse de vitesse en $km.h^{-1}$ par rapport à la vitesse moyenne qui est de $53km.h^{-1}$. On remarquera que pour certains points entourés en vert, la congestion fait bien son apparition sans que $\bar{v}^{T,500,t}$ ne soit impactée. Certes le système se congestionne et certains trains prennent du retard mais ces trains peuvent compenser ce retard un peu plus tard en roulant plus vite. C'est l'état du trafic de congestion résiliente.

la congestion apparaît, la chute de vitesse (en $km.h^{-1}$) par rapport à la vitesse planifiée. Une chute de $0km.h^{-1}$ indique que si la congestion a bien lieu, les trains sont capables de compenser entièrement les chutes de vitesse qu'ils subissent sur quelques cantons en roulant plus vite sur les autres cantons.

Grâce à cet indicateur on peut désormais distinguer les même trois états de trafic que dans la partie 2.5.2. Rappelons quels sont ces trois états de trafic :

- L'état fluide : la congestion n'apparaît pas.
- L'état congestionné résilient : la congestion apparaît mais n'entraîne pas une chute de la vitesse moyenne des trains. Des petits retards apparaissent mais chaque train qui prend du retard est ensuite capable de le rattraper.
- L'état congestionné dégradé : la congestion apparaît et entraîne une chute de la vitesse moyenne des trains, les retards s'accumulent.

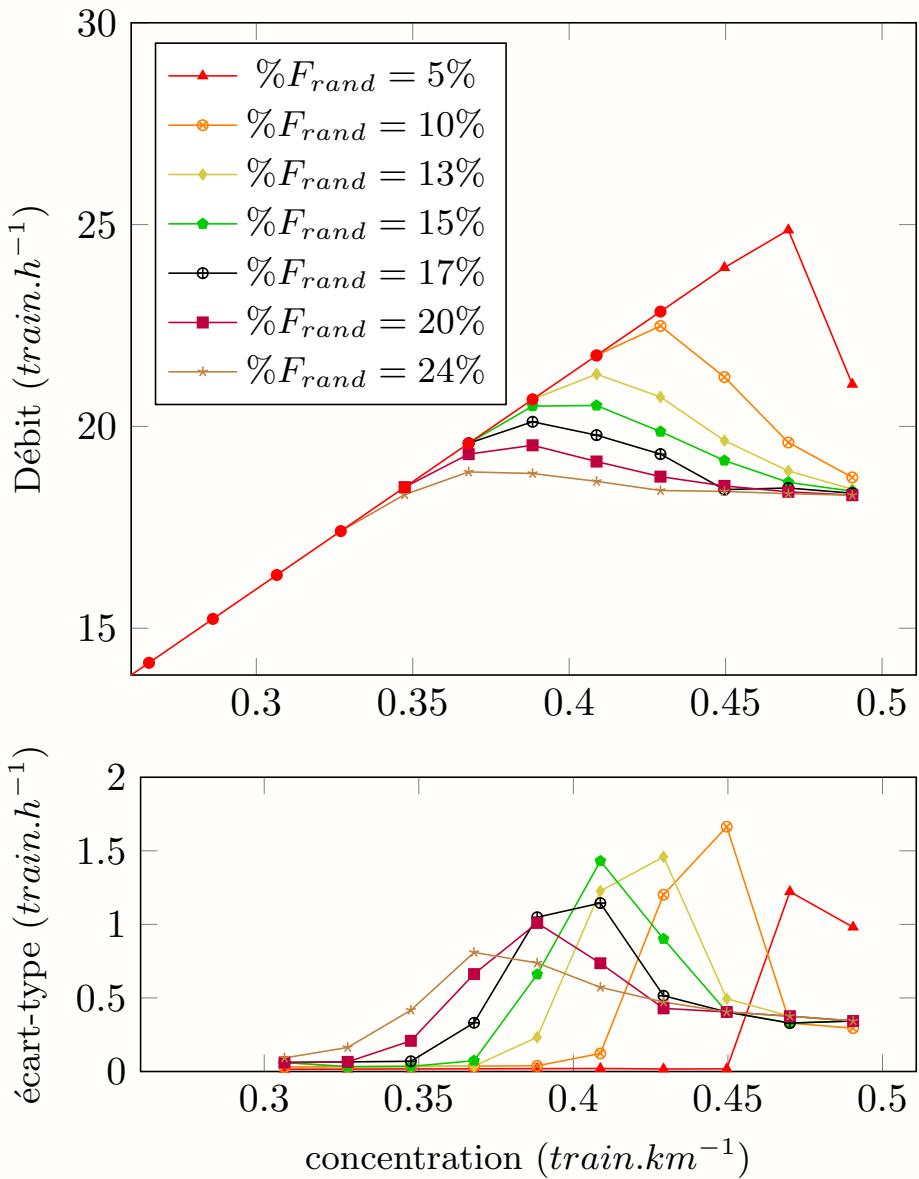
3.3.3 Diagramme Fondamental de Ligne Ferroviaire

Le DFLF représente le débit moyen d'une ligne ou d'une portion de ligne ferroviaire en fonction de la concentration de trains sur cette même ligne ou portion. La figure 3.8 représente un DFLF dans lequel le hasard est introduit à l'aide de la force aléatoire. La figure 3.9 représente un DFLF dans lequel le hasard est introduit à l'aide du temps de stationnement. Les points de ce diagramme ont été obtenus de la manière suivante :

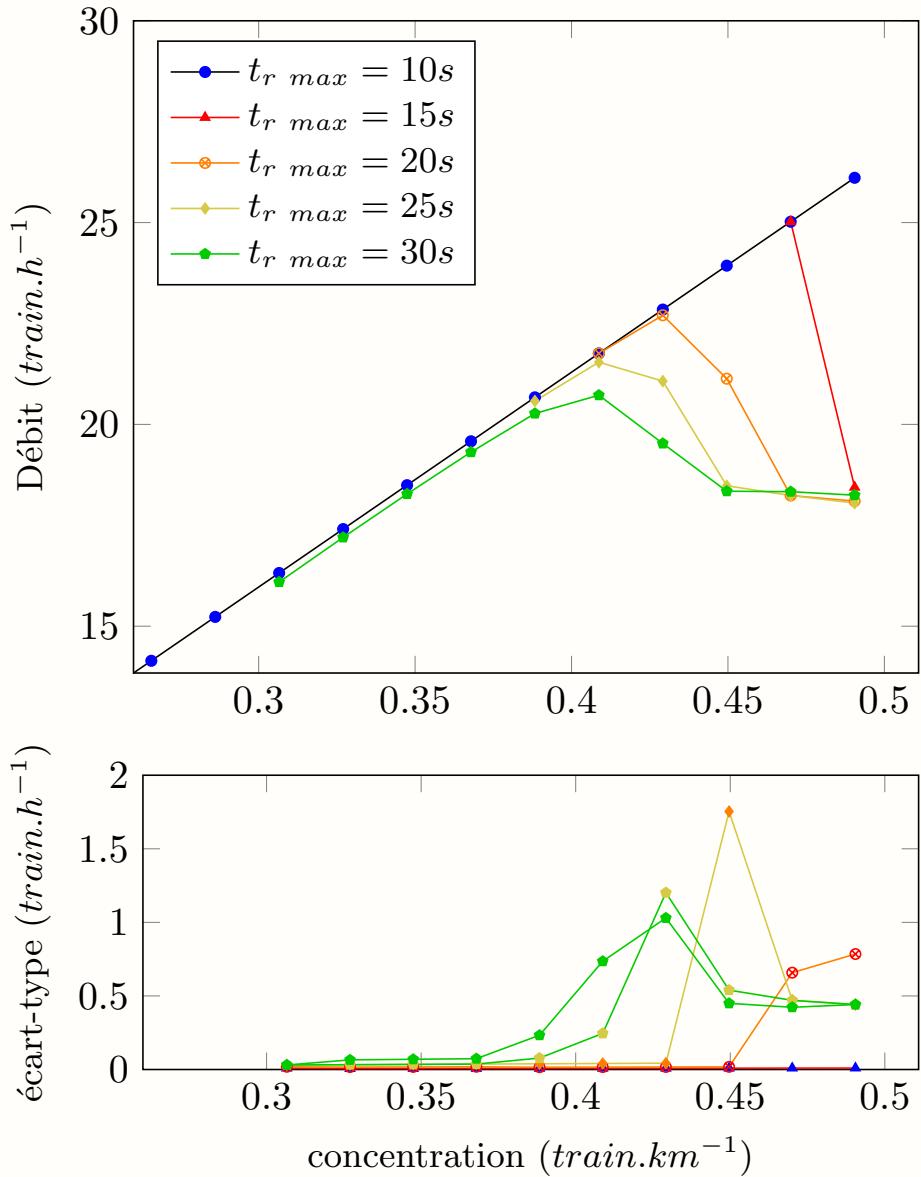
- Pour $v_{max} = 90km.h^{-1}$ et pour une intensité de la force aléatoire, du temps additionnel de stationnement et du nombre de trains variant respectivement de 5% à 24%, de 10s à 30s et de 1 à 24, on a répété les simulations 500 fois dans le cadre de la méthode de Monte-Carlo.
- Pour chaque simulation le débit moyen est obtenu sur l'ensemble de la ligne simulée et sur la fenêtre temporelle des trente dernières minutes des simulations.
- Pour chaque valeur du nombre de trains on obtient une concentration et une distribution des débits moyens, la moyenne de cette distribution et son écart type sont représentés sur les figures 3.8 et 3.9.

Tant que les retards ne s'accumulent pas, c'est-à-dire tant que le système ne passe pas dans un état congestionné dégradé, le plan de transport initialement prévu peut être conservé. Les petites perturbations que subit ce plan de transport ne modifient pas le débit des trains et tous les trains prévus peuvent circuler. En revanche, lorsque le système se trouve dans l'état congestionné dégradé, le débit devient inférieur au débit planifié et on ne peut faire circuler le trafic désiré.

On remarque par ailleurs sur la figure 3.9 que pour $t_{r max} = 30s$, le débit observé est inférieur au débit prévu par le plan de transport. Et ce, même en dehors de la zone habituelle de congestion. Pour $t_{r max} = 25s$, le débit réel est également plus faible que le débit planifié mais dans des proportions moindres. Il ne s'agit pas à proprement parler d'un phénomène de congestion car les trains n'interagissent pas entre eux. C'est simplement

Figure 3.8 - DFLF avec gares, variations de la force aléatoire

Ce diagramme fondamental de ligne ferroviaire représente le débit ferroviaire en fonction de la concentration pour $v_{max} = 90 \text{ km} \cdot \text{h}^{-1}$ et pour une intensité de la force aléatoire et du nombre de trains variant respectivement entre 5% et 24% d'une part et entre 1 et 24 d'autre part. Ce débit ferroviaire est estimé à partir de la vitesse moyenne $\bar{v}^{T,500,t}$, on trace également l'écart type associé (celui de la distribution des $\bar{v}^{T,t}$).

Figure 3.9 - DFLF avec gares, variations du temps d'arrêt

Ce diagramme fondamental de ligne ferroviaire représente le débit ferroviaire en fonction de la concentration pour $v_{max} = 90 \text{ km.h}^{-1}$ et pour un temps additionnel de stationnement maximum et du nombre de trains variant respectivement entre 10s et 30s d'une part et entre 1 et 24 d'autre part. Ce débit ferroviaire est estimé à partir de la vitesse moyenne $\bar{v}^{T,500,t}$, on trace également l'écart type associé (celui de la distribution des $\bar{v}^{T,t}$).

un effet que l'on doit à un ingrédient de la construction du temps d'arrêt additionnel qui n'est pas présent dans la construction de la force aléatoire. Selon la nature de l'aléa que l'on introduit, voilà comment le système se comporte :

- Imaginons qu'un train unique circule sur notre ligne, il ne peut donc pas y avoir de congestion. La force aléatoire puisqu'elle est centrée, n'introduit pas une asymétrie dans les écarts à l'horaire qu'elle génère. Le système reste globalement à l'heure.
- En revanche, dans le cas où il y a plusieurs trains, à mesure que ces aléas augmentent, certains trains se rapprochent ce qui crée une situation de congestion. C'est alors la force d'interaction, qui elle est forcément asymétrique qui génère du retard.
- Imaginons à nouveau qu'un train unique circule sur notre ligne. Le temps d'arrêt additionnel étant construit de manière asymétrique, il ne peut que retarder les trains. Ce retard est rattrapé grâce à la force horaire entre deux gares. Mais à mesure que l'on augmente l'amplitude de l'aléa, le système a de plus, en plus de mal à rattraper ce retard. C'est ce qui arrive pour $t_{r \ max} = 25s$ et qui devient suffisamment important pour que l'on arrive à l'observer sur la figure 3.9 pour $t_{r \ max} = 30s$.

Une vision de l'esprit qui peut être efficace pour mieux appréhender ce phénomène, peut être obtenue en décomposant un aléa non centré en deux composantes. Par exemple, un temps additionnel choisi au hasard de manière uniforme entre 0s et 30s peut très bien être décomposé en un temps additionnel fixe de 15s et un temps additionnel choisi au hasard de manière uniforme entre -15s et 15s, c'est-à-dire un aléa centré. D'après ce qu'on vient de dire, une baisse du débit peut ne pas être due à un phénomène de congestion. En effet, cette baisse peut apparaître alors qu'il n'y a qu'un seul train et que la congestion implique justement plusieurs trains. Les sources d'aléas non centrées vont ralentir les trains⁶, ce ralentissement peut être compensé, par les marges de vitesses par exemple, tant qu'il reste suffisamment faible. Au-delà d'un seuil de dé-centrage, la vitesse moyenne baisse.

3.4 Conclusions sur le modèle développé

On a, dans ce chapitre mais aussi dans le chapitre 2 abordé notre modèle sous deux angles différents. L'aspect statique du modèle concerne essentiellement les DFLF tracés à partir du modèle pour lesquels la dynamique est forcément absorbée par le calcul d'une moyenne sur des durées de 30 min. L'utilisation des DFLF ne permet donc pas d'appréhender, ni la dynamique de l'apparition de la congestion, ni la dynamique de la congestion elle-même. Pourtant, ces aspects dynamiques sont bien réels. Ces aspects ont été principalement étudiés au travers de la probabilité d'apparition de la congestion mais aussi par l'analyse de l'évolution de la vitesse moyenne des trains au cours du temps. Cette dernière partie est l'occasion de faire un bilan des principaux résultats obtenus avec notre modèle.

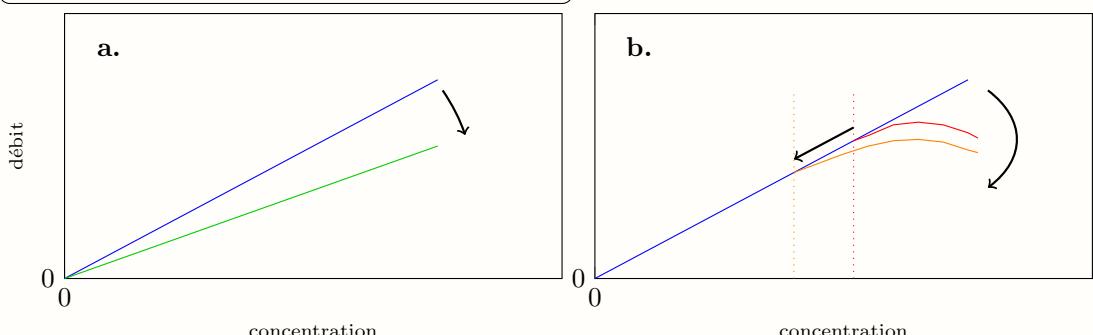
3.4.1 Aspects statiques

Sur la figure 3.10 on a utilisé la décomposition d'un aléa, déjà présentée en 3.2.2, en une partie fixe et un aléa centré. La sous-figure a représente les effets sur le DFLF d'un

6. le cas où elles accélèrent les trains n'est pas intéressant.

temps additionnel déterministe (et dans le cas présenté ce temps est un retard). À mesure que ce temps additionnel augmente, on observe tout d'abord une résilience du système due aux marges de rattrapage, puis une baisse de la vitesse moyenne qui ne dépend pas de la concentration. La sous-figure **b** représente les effets sur le DFLF d'un aléa centré. Plus cet aléa est grand, plus la chute de débit apparaît à de faibles concentrations. Concernant le premier cas, il ne s'agit pas à proprement parler d'un phénomène de congestion au sens de la définition présentée en 2.3. Ce phénomène se manifeste en effet sans qu'il y ait une interaction entre les trains. Le second cas est, au contraire, un phénomène de congestion qui se manifeste avec interaction entre les trains. Ces résultats sont repris à la toute fin du chapitre 4.

Figure 3.10 - Effets des aléas sur le DFLF



Effets des aléas sur le DFLF. La sous-figure **a** représente les effets sur le DFLF d'une perturbation déterministe : un changement de la pente du diagramme. La sous-figure **b** représente les effets sur le DFLF de la composante aléatoire centrée d'un aléa : un effondrement du débit pour les hautes concentrations.

3.4.2 Aspects dynamiques

On a pu observer les formes que prenaient les probabilités d'apparition de la congestion dans notre modèle. On a alors construit un modèle macroscopique explicatif qui reproduit les courbes de probabilités d'apparition de la congestion. Ce modèle qui repose sur les processus de Markov, s'est montré efficace pour reproduire les courbes d'apparition de la congestion tant pour le modèle avec gares que pour le modèle sans gare. Cependant dans un cas comme dans l'autre nous avons testé trop peu de configurations pour affirmer que ce processus de Markov s'applique effectivement dans tous les cas possibles.

Sur la modélisation avec gares, nous avons testé une implémentation des aléas, d'une part en utilisant une force aléatoire, d'autre part en utilisant un temps d'arrêt supplémentaire aléatoire. Dans le premier cas les aléas s'expriment sur tous les cantons⁷. En revanche, dans le second cas, les aléas ne s'expriment qu'en gare. Cette différence a une conséquence sur la dynamique d'apparition de la congestion. Comme on l'a montré dans

7. Certes la force s'exprime moins fortement sur les cantons *c* et *g*.

la partie 3.2.2, les probabilités d'apparition de la congestion ont tendance à augmenter moins vite de zéro au début des simulations si la situation n'est pas homogène. Il ne faut surtout pas en conclure que ces situations sont avantageuses car le passage à une situation hétérogène engendre également une chute de débit comme on l'a vu en 3.2.2 avec l'exemple d'un motif *abbcg*.

Chapitre 4

Confrontation avec les données

Sommaire

4.1	Données	110
4.2	Impossibilité d'un diagramme fondamental local	112
4.3	Vers un diagramme fondamental de ligne ferroviaire	113
4.3.1	Principes du diagramme fondamental de réseau en routier	113
4.3.2	Méthode de construction du DFLF	114
4.3.3	Résultats et analyses	115
4.4	Confrontation des données avec le modèle	119
4.4.1	Diagramme fondamental de ligne ferroviaire moyené : cas réel	119
4.4.2	Mise en perspective avec la modélisation	120
4.5	Discussion et conclusion	122

Introduction

Jusqu'ici, les travaux présentés étaient purement théoriques. Afin de conforter l'intérêt que pourrait représenter les résultats issus du modèle développé dans les chapitres 2 et 3, mais aussi les hypothèses de modélisation que l'on a faite, une confrontation critique avec des données expérimentales est inévitable. Le Diagramme Fondamental de ligne Ferroviaire (DFLF), qui est construit à partir de données empiriques dans ce chapitre, va nous servir d'élément de comparaison.

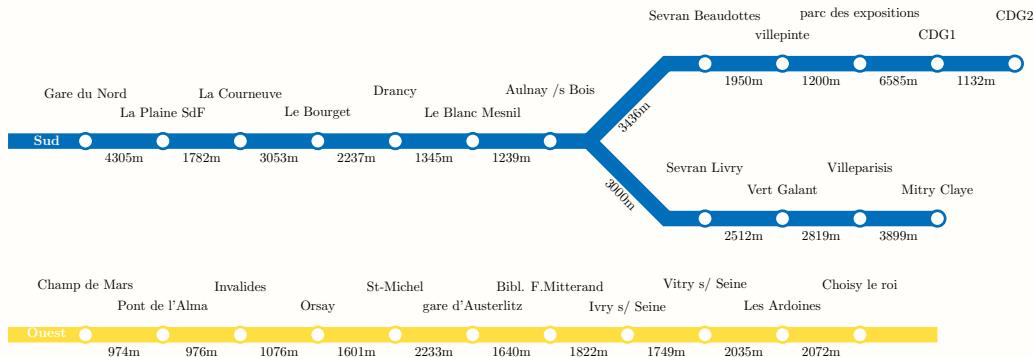
Dans ce chapitre, nous allons tout d'abord présenter brièvement les tronçons de ligne du réseau ferroviaire sur lesquelles nous avons travaillé, ainsi que les données associées. À partir de ces données nous montrerons ensuite les écueils qui s'opposent à la construction et à l'utilisation d'un DF local pour le ferroviaire. Analyser la congestion en ferroviaire impose donc une autre approche, nous allons donc présenter ici, à travers la bibliographie, l'existence d'un diagramme non local en routier qui utilise les données collectées sur l'ensemble d'un réseau routier. C'est le diagramme fondamental de réseau (DFR). En s'inspirant de ce diagramme nous allons construire un diagramme non local pour le ferroviaire en utilisant les données disponibles sur une portion de ligne. C'est le diagramme fondamental de ligne ferroviaire (DFLF). Nous proposons deux versions de ce diagramme. La première version utilise les données d'une seule journée de trafic à la fois, cette version permet de mettre en évidence visuellement des phénomènes de dégradation du trafic. La seconde version utilise les données de toutes les journées disponibles dans un seul diagramme. Cette version correspond plus ou moins au DFLF présenté à la fin des chapitres 2 et 3 ce qui nous permet de comparer les résultats obtenus de manière empirique avec ceux, obtenus de manière théorique, des chapitres précédents.

4.1 Données

La région parisienne, fortement peuplée, a une demande de transport qui se situe à un niveau très élevé. Elle connaît dans ce contexte, des problèmes de transport en commun. Elle est desservie par deux réseaux de chemin de fer : le métro et le réseau express régional (RER). Ce dernier transporte environ 7 millions de passagers par jour sur 14 lignes. Pour nos travaux, nous choisissons de nous concentrer sur la partie RER exploitée par la SNCF. Aujourd'hui, la demande croissante rend le RER sur-saturé. Nous avons ici choisi d'étudier la partie nord de la ligne B et la partie centrale de la ligne C. La raison du choix de ces deux lignes est à la fois due à un trafic ferroviaire très élevé et à un niveau préoccupant de retards sur ces deux lignes. La figure 4.1 représente des deux domaines analysés : le nord de la ligne B entre Paris Gare du Nord, CDG Aeroport terminal 2 et Mitry Claye, et la section centrale de la ligne C entre Champ de Mars et Choisy le Roi.

Nous allons maintenant présenter les données dont nous disposons pour ces segments de lignes. De même que des capteurs sont régulièrement enterrées sous la chaussée pour analyser le trafic routier, des capteurs sont présents régulièrement sur les voies de chemins de fer du réseau Transilien. Ces capteurs que l'on appelle des «balises de détection des trains» ou encore des «localiseurs» permettent de détecter le passage des trains et

Figure 4.1 - Portion du réseau analysé



Portion du réseau analysé : Le nord de la ligne B en bleu soit 40km de ligne la portion centrale de la ligne C en jaune qui fait 16km. Le nom de stations et la distance entre elles est également indiqué.

d'enregistrer l'heure de cet événement. À l'inverse, les capteurs routiers enregistrent souvent en plus de son heure de passage, la vitesse d'un véhicule. Les caractéristiques de ces balises ferroviaires sont :

- Une position sur une voie ;
- Un sens de circulation ;
- Elles sont placées à la sortie d'un canton.

Pour mener à bien notre étude, il faut exploiter une base de données qui rassemble toutes les informations collectées par ces balises. Comme on l'a vu en 1.3.6, en trafic ferroviaire, la planification relègue à l'arrière-plan les concepts d'offre et de demande utilisés en routier au profit des concepts de graphique des horaires planifiés ou plan de transport et de graphique des horaires réalisés. On utilise donc une seconde base de données qui nous fournit non plus les données collectées par les balises mais celles qui aurait dû l'être si le plan de transport avait été rigoureusement respecté.

Dans le cas où les voies sont dédoublées comme cela peut être le cas à l'entrée des gares, la base de données utilisée agrège les données et ne permet pas de distinguer la voie utilisée. Une fois les données regroupées de cette manière, la densité des balises sur le réseau correspond en règle générale à une balise par gare et c'est effectivement le cas pour les zones étudiées. Comme chacune des deux lignes présente deux directions, on compte pour chaque jour quatre ensembles de données collectées. Les jours avec forte baisse de la demande (comme les week-ends et jours fériés) ou pour lesquels un événement spécial (travaux prévus, match de football ...) impacte la circulation sont rejettés. On a également procédé à un nettoyage de la base de données qui peut comporter quelques erreurs d'origine humaine. En fin de compte, 32 jours de données sont disponibles pour la ligne B et 34 pour la ligne C. La liste des jours utilisés pour chaque ligne est présentée dans un calendrier figure 4.2.

Figure 4.2 - Calendrier des jours analysés

Calendrier des jours analysés. Les jours en rouge-gras concernent les lines B et C ceux en vert-italique ne concernent que la ligne C.

4.2 Impossibilité d'un diagramme fondamental local

On vient de dire que nos balises ne mesuraient pas la vitesse des trains mais enregistraient seulement leur heure de passage. Or pour construire un DF, il nous faut mesurer localement deux des trois variables suivantes : la vitesse, la concentration et le débit (voir la partie 1.3 et en particulier la formule (1.7)). Ici la seule variable mesurée est le débit local. Les données disponibles ne nous permettent donc pas, dans ce cas, de tracer un DF. Pour pourvoir mesurer une vitesse, il nous faut utiliser un second temps de passage que l'on peut récupérer sur une seconde balise.

Si on construit un diagramme en utilisant les données d'une seconde balise, il perd son caractère local. Néanmoins, un tel diagramme est, au vu des données disponibles ce que l'on peut tracer de plus proche d'un DF.

Pour construire un tel diagramme, on utilise une zone d'Édie (voir 1.3) qui a un pas spatial de la zone entre les deux balises et un pas temporel de 10min . Un train qui ne serait présent sur la portion étudiée qu'une fraction de ces 10 min ne sera comptabilisé que comme une fraction de train. Toutes les minutes, on évalue avec la méthode d'Édie, le débit et la concentration. On présente un des diagrammes ainsi obtenu sur la figure 4.3. Ce diagramme a été construit avec les données de la journée du 21 mai 2013 pour la ligne B entre les stations Aulnay sous Bois et le Blanc Mesnil. Nous avons augmenté le temps d'agrégation des données sans parvenir à réduire pour autant la dispersion du diagramme. Nous ne sommes pas tout à fait à l'échelle locale, les vitesses étant moyennées sur une distance de l'ordre du kilomètre, ce qui contribue à diminuer la dispersion que l'on aurait à l'échelle locale. La dispersion des points sur ce diagramme est d'ailleurs suffisante pour

nous conduire à penser qu'à l'échelle locale le trafic ferroviaire ne peut être considéré avec des variables continues. Un DF (local) n'est pas un outil utile pour le ferroviaire. On va maintenant présenter le DFR en routier.

Figure 4.3 - Diagramme fondamental expérimental entre deux balises

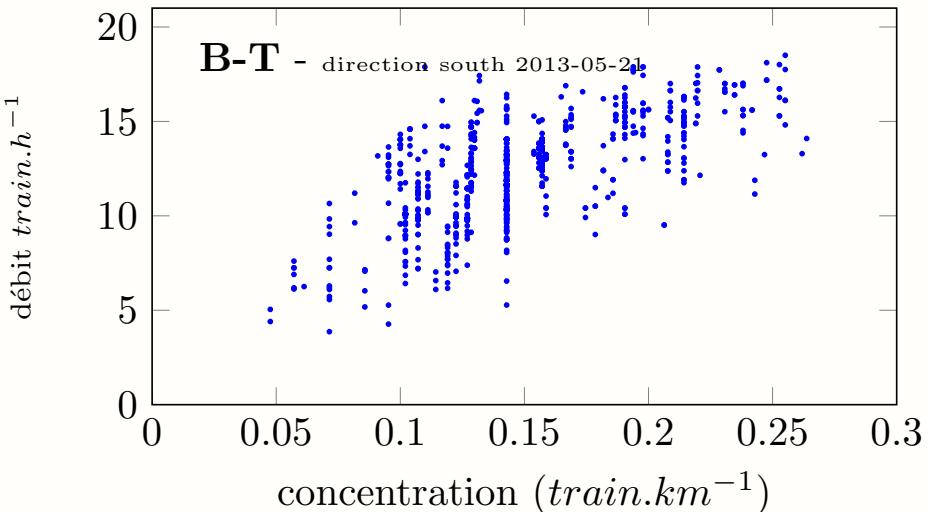


Diagramme fondamental expérimental entre deux balises

4.3 Vers un diagramme fondamental de ligne ferroviaire¹

4.3.1 Principes du diagramme fondamental de réseau en routier

Depuis plus de cinquante ans, les détecteurs régulièrement disposés sous la chaussée des routes permettent de détecter le passage des véhicules et d'enregistrer leur vitesse. Ces détecteurs, des bobines de cuivre, voient leur impédance varier lorsqu'une masse riche en fer, en l'occurrence un véhicule, passe au-dessus. Si une boucle suffit à détecter le passage d'un véhicule, ces boucles sont souvent installées par paires sous la chaussée afin d'évaluer également la vitesse du véhicule détecté. À partir de ces observations, on peut évaluer des variables caractérisant la circulation vue dans son ensemble mais toujours à l'échelle locale (à l'endroit du détecteur). Ces variables sont la vitesse moyenne et le débit moyen à l'échelle locale. La densité (nombre de véhicules pour une longueur de tronçon donnée) peut également être extraite par une approximation grossière en mesurant la fraction de temps durant laquelle la boucle est occupée par les véhicules.

Si l'on dispose assez de boucles sur l'ensemble d'un réseau, il est possible d'évaluer ces grandeurs non plus de manière locale mais de manière globale, c'est-à-dire moyennée sur l'ensemble du réseau. Un tel calcul n'est pas une idée récente : après la publication initiale de cette idée dans GODFREY [1969], Herman et Prigogine proposent dans HERMAN

1. cette partie est un résumé succinct de l'article CUNIASSE et collab. [2015]

et PRIGOGINE [1979], d'examiner la relation entre le débit moyen sur un réseau routier et la concentration moyenne sur ce même réseau. Ces articles ne se sont consacrés qu'à la partie fluide du diagramme où les débits moyens augmentent avec la concentration. MAHMASSANI et collab. [1984] évoquent la possibilité d'événements aléatoires de blocages des voies dans un réseau urbain de type grille rectangulaire simulé avec NETSIM. Plus le rapport entre le temps de blocage des voies et le temps de simulation total est grand, plus la pente de la relation entre le débit moyen et la concentration moyenne est faible. Les auteurs mettent en évidence que la vitesse des véhicules dans le système considéré s'obtient bien en faisant le rapport entre la concentration moyenne et le débit moyen. Même avec ces blocages de voies, il est à noter que la partie congestionnée du diagramme fondamental obtenu n'est pas observée.

Dans DAGANZO [2007] Daganzo suppose, par analogie avec le DF local, qu'au-dessus d'un seuil donné pour la concentration moyenne (valeur critique de la concentration), une augmentation du nombre de véhicules conduit à une diminution du débit total. Cela a été observé expérimentalement pour la première fois par Geroliminis et Daganzo un an après DAGANZO et GEROLIMINIS [2008]; GEROLIMINIS et DAGANZO [2007, 2008]; GEROLIMINIS et SUN [2011] pour le centre congestionné de la ville de Yokohama. Cette relation entre le débit total et l'accumulation totale moyenne pendant les périodes de typiquement quelques minutes sur un réseau a d'abord été appelée par Daganzo et Geroliminis : un «diagramme fondamental macroscopique» (appelé MFD² dans de nombreux documents). En accord avec les auteurs de SABERI et MAHMASSANI [2012]; SABERI et collab. [2014], nous utiliserons le terme «diagramme fondamental de réseau» DFR ; une dénomination définitive ne faisant pas encore l'objet d'un consensus dans la communauté scientifique.

Nous nous sommes inspirés de ce diagramme non local construit pour le trafic routier pour développer un outil similaire pour le trafic ferroviaire le DFLF. Nous allons maintenant présenter la méthode de construction d'une première version de ce diagramme qui regroupe les données d'une seule journée.

4.3.2 Méthode de construction du diagramme fondamental de ligne ferroviaire

On appelle le passage d'un train sur une balise un événement. Pour chaque jour et pour chaque ensemble de données, chaque événement de chaque balise est pondéré par la moitié de la somme des deux distances qui la séparent des deux balises voisines. On s'affranchit ainsi du fait que l'espacement entre deux balises consécutives n'est pas constant. Notre méthode permet donc d'éviter de sur-représenter les zones où les balises sont plus rapprochées.

Sur l'ensemble de la portion étudiée, on compte à chaque instant le nombre de trains présents. Ce nombre, une fois divisé par la longueur de la portion de ligne considérée, nous donne la concentration moyenne à chaque instant. On calcule ensuite une moyenne tem-

2. sigle en anglais

porelle de cette concentration instantanée grâce à une fenêtre temporelle de 10 min et ce, toutes les minutes.

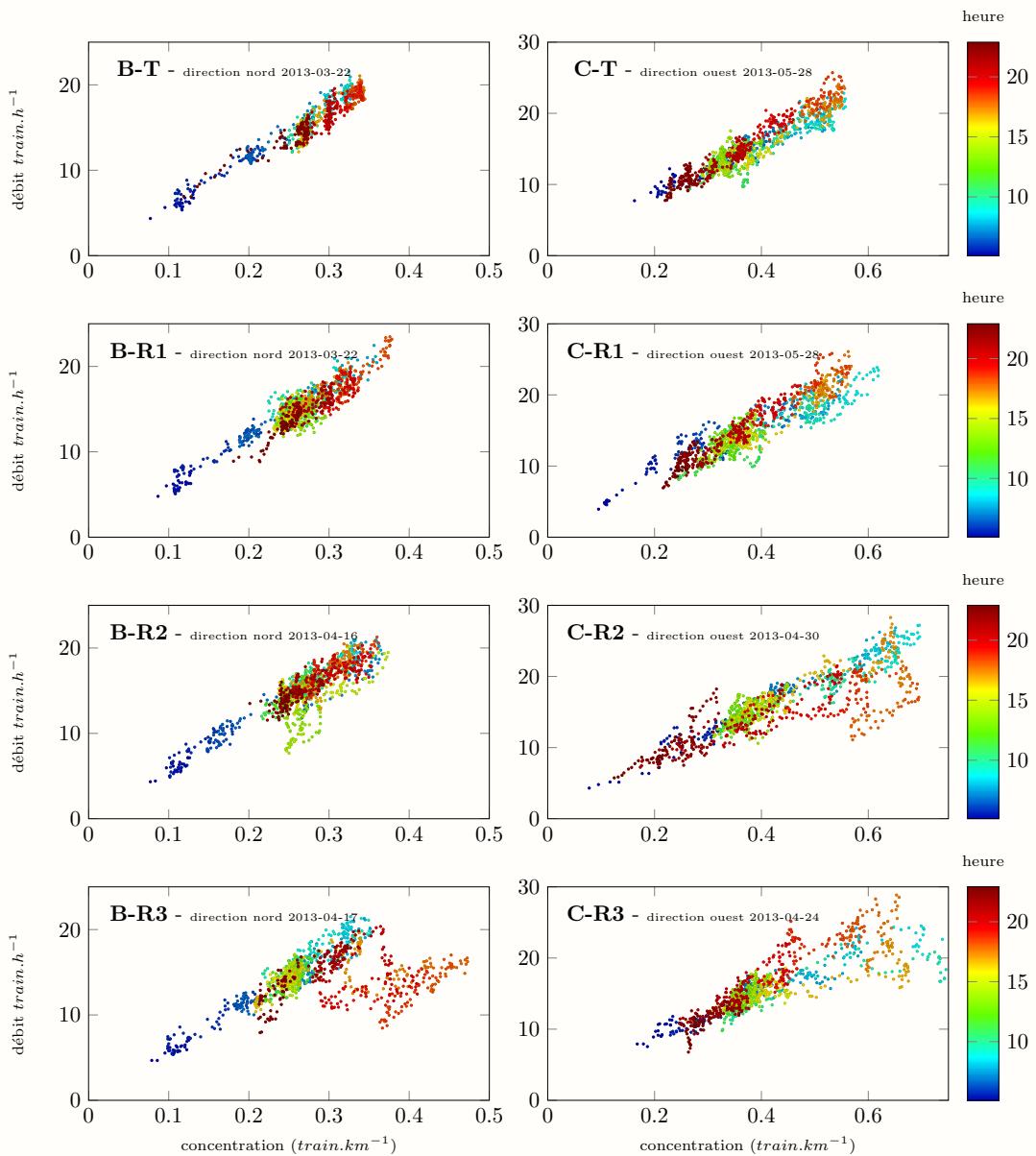
Nous sommes en train de construire une zone d'Édie (voir 1.3) qui a un pas spatial de la taille de la zone étudiée et un pas temporel de 10min. Un train qui ne serait présent sur la portion étudiée qu'une fraction de ces 10 min ne sera comptabilisé que comme une fraction de train. On compte également le nombre total d'événements relevé par des balises sur cette même portion de réseau par fenêtre temporelle de 10 min et également toutes les minutes. Chaque événement associé à une balise étant pondéré tel que décrit ci-dessus. La valeur obtenue est la vitesse moyenne de l'ensemble de trains présents sur la portion étudiée pendant 10 min. Le débit moyen correspondant peut être obtenu en multipliant cette vitesse moyenne avec la concentration moyenne. Nous traçons enfin le débit en $train.h^{-1}$ en fonction de la concentration en $veh.km^{-1}$. Nous appelons le diagramme obtenu un Diagramme Fondamental de Ligne Ferroviaire ou DFLF.

4.3.3 Résultats et analyses

La figure 4.4 présente le DFLF pour 2 lignes : la ligne B - direction nord (colonne de gauche) et la ligne C - direction ouest (colonne de droite). Pour faciliter la localisation temporelle des événements observés, un code couleur indique l'heure. Les deux schémas **B-T** et **C-T** en haut de la figure 4.4 représentent la situation prévue (T signifiant théorique ou planifié) sur une journée typique. Les schémas qui viennent en dessous représentent le diagramme obtenu pour quelques jours sur des données réelles. Selon le jour, on observe une dispersion plus ou moins grande du diagramme. Comme en routier, cette dispersion est d'autant plus grande que la largeur de la fenêtre temporelle choisie est petite, notre choix de fixer cette largeur à 10 min permet de limiter cette dispersion. En contrepartie, les évolutions du diagramme à des échelles de temps plus petites que 10 min sont filtrées. Dans tous les cas, on observe que la plupart des points semblent alignés le long d'une droite. Cet alignement sur une droite est d'autant plus visible sur les diagrammes correspondant aux horaires théoriques.

On vient de voir avec la figure 4.3 que le tracé d'un DF local en ferroviaire n'est pas une bonne idée car les variables du débit et de la concentration ne variaient pas de manière continue ce qui conduisait à une dispersion extrême des points du DF. En agrégant les mesures sur une zone suffisamment large et sur une fenêtre temporelle suffisamment grande, cette dispersion diminue de façon significative au point que les DFLF théoriques prennent, comme on vient de le constater, la forme d'une droite passant par zéro.

Or il se trouve qu'observer cette droite a du sens : en effet le plan de transport est conçu de telle sorte que les véhicules puissent rouler à voie libre. Dans ces conditions (certes idéales) la vitesse moyenne d'un véhicule donné sur un parcours donné ne dépend pas du nombre de véhicules présents sur ce parcours. On s'attend donc bien à une relation de proportionnalité entre le débit et la concentration c'est-à-dire à une droite qui passe par l'origine. On a donc utilisé les DFLF correspondants aux horaires théoriques et choisi comme critère, l'observation d'une droite acceptable pour considérer que les fenêtres spa-

Figure 4.4 - Exemples de DFLF


Exemples de DFLF. **B-T & C-T** sont les diagrammes théoriques ou planifiés, **B-R1, B-R2, B-R3, C-R1, C-R2 et C-R3** sont des exemples de diagrammes réalisés sur quelques journées.

tiales et temporelles étaient suffisantes. Ainsi calculées, on peut observer les concentrations moyennes et les vitesses moyennes comme des variables à peu près continues.

Le choix d'une fenêtre temporelle de 10 min pour le calcul des moyennes repose sur un compromis entre deux impératifs : l'obtention d'une droite acceptable mais aussi d'une fenêtre temporelle pas trop grande pour «flouter» le moins possible le déroulement des événements. Mais il faut noter que ce compromis n'a été tranché que visuellement.

Pour vérifier à posteriori l'hypothèse d'une forme de droite passant par l'origine, nous allons procéder à des régressions. Il y a une bonne raison pour la courbe de passer par le point (0, 0). En effet, zéro train implique forcément zéro mesure sur les balises. On ne teste donc que des modèles de régression qui ont une ordonnée nulle à l'origine. Finalement, nous testons deux propositions de régressions y_1 et y_2 que nous testons sur les DFLF théoriques uniquement :

$$y_1 = \beta_1 \cdot x$$

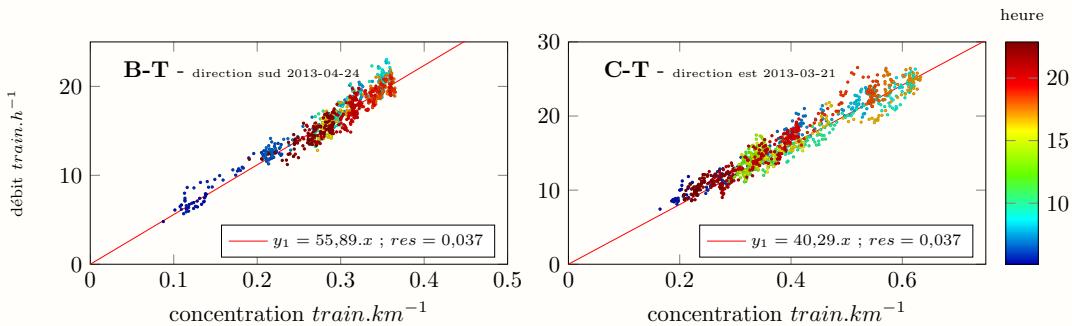
$$y_2 = \beta_2 \cdot x + \gamma \cdot x^2$$

Comme le second modèle de régression y_2 n'est pas linéaire, on utilise le résidu χ^2 pour comparer les deux régressions plutôt que le coefficient de détermination R. Le tableau 4.1 présente la valeur moyenne de χ^2 pour chaque cas. La régression y_2 a été rejetée parce qu'elle ne donne pas des résultats significativement meilleurs que la régression y_1 et qu'elle a besoin d'un paramètre en plus. Selon le principe de parcimonie, on choisit donc de ne garder qu' y_1 . Certains exemples de régressions y_1 sont représentés dans la figure 4.5.

Tableau 4.1 - Moyennes des résidus des régressions

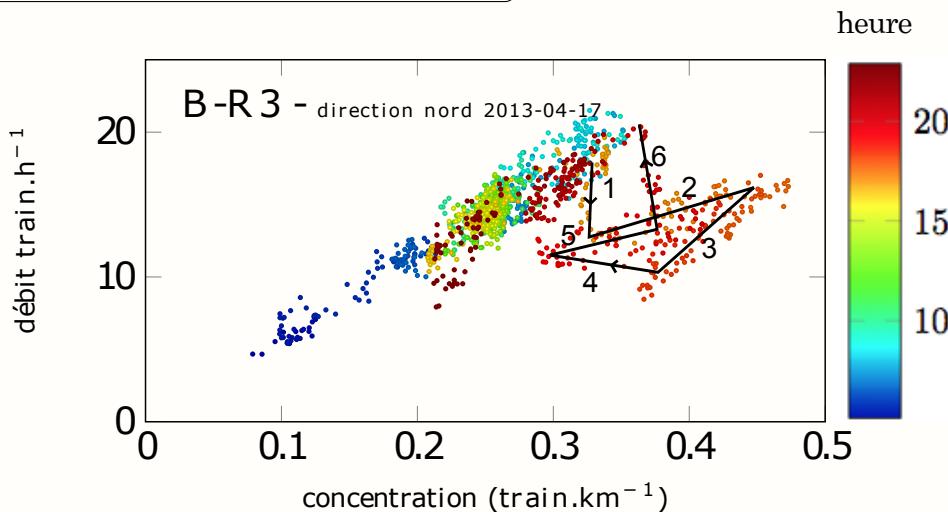
Ligne et direction	B-sud-T	B-nord-T	C-est-T	C-ouest-T
Nbr jours	15	15	14	14
y_1	0.0378	0.0379	0.0366	0.0344
y_2	0.0379	0.0377	0.0362	0.0339

Nous pouvons voir sur la figure 4.4 qu'en plus de la droite que l'on vient d'évoquer, diverses formes sont observées pour les diagrammes qui correspondent à des situations réelles. Ces formes placées en dessous de la droite théorique correspondent à des chutes du débit. Nous avons retenu l'un de ces diagrammes pour lequel nous avons regardé précisément l'évolution du trafic sur la période qui correspond à cette chute de débit. Pour cette analyse, on reprend la sous-figure B-R3 de la figure 4.4, et on trace sur la figure 4.6 le déroulement des événements suivants : la baisse de débit a lieu juste avant l'heure de pointe à 16h20 lorsque deux trains consécutifs entrent sur le secteur étudié avec un retard de 12 minutes forçant deux autres trains à ralentir dans la zone étudiée. À partir de ce moment et jusqu'à 20h30 cela entraîne des ralentissements qui sont régulièrement observée

Figure 4.5 - Exemples de DFLF avec interpolation


Exemples de DFLF avec interpolé par une relation simple : $y_1 = \beta_1 \cdot x$. Les sous-figures **B-T** et **C-T** correspondent aux LFDF théoriques pour la ligne B dans la direction sud et la ligne C dans la direction est. L'unité de la pente des interpolations est une vitesse : $\beta = [km.h^{-1}]$

dans la zone étudiée, et qui maintiennent un niveau de débit faible. Six phases peuvent être identifiées au cours de cette chute du débit. Cette décomposition dépend de la lecture plus ou moins fine que l'on peut avoir des événements.

Figure 4.6 - Analyse d'une chute de débit


Analyse d'une chute de débit sur l'exemple du DFLF **B-R3**, ligne B dans la direction nord le 2013-04-17. On distingue 6 phases en s'aidant des données détaillées du trafic.

1. **16 :20** Les premiers ralentissements apparaissent entraînant une chute du débit.
2. **17 :20-19 :00** C'est l'heure de pointe, le nombre de trains augmente sur la portion étudiée.
3. **19 :10-19 :50** Ralentissement sévère pour 3 trains qui entraîne une chute encore plus importante du débit.
4. **19 :20-20 :00** Plusieurs trains sont supprimés, le nombre de trains sur le réseau diminue.

5. **20 :00-20 :30** Fin des suppressions de trains, le nombre de trains augmente à nouveau.
6. **20 :30** Les vitesses redeviennent normales, fin de la chute de débit, un retard de 15 min subsiste.

La comparaison entre le routier et le ferroviaire en [1.3.6](#) a mis en évidence un certain nombre de différences entre ces deux modes de transport. Nous venons de montrer qu'en tenant compte de ces différences, les outils du trafic routier peuvent être adaptés pour étudier le trafic ferroviaire. Nous allons maintenant construire une variante du DFLF plus adaptée pour une confrontation avec nos résultats théoriques

4.4 Confrontation des données avec le modèle

Le fait que certains DFLF réalisés présentent des formes de droite et d'autres des formes plus complexes pourrait nous inciter à classer ces diagrammes en deux catégories mais il existe une vision plus adaptée pour comprendre la diversité de ces diagrammes comme issus d'un même processus. En effet, le fait que ces diagrammes présentent des formes qui ne semblent pas se répéter est un bon argument pour justifier le fait que l'on ait modélisé les perturbations du trafic ferroviaire comme des variables aléatoires dans les chapitres [2](#) et [3](#). Il convient donc à présent de confronter les résultats obtenus à partir de notre étude de ligne réelle avec les résultats issus du modèle théorique que l'on a construit dans ces chapitres. Nottons qu'en l'absence de calibration des simulations réalisées au chapitre [3](#), cette comparaison ne peut être que qualitative pour l'instant.

L'obtention des DFLF à partir de la modélisation dans les chapitres précédents reposait entre autre sur l'usage de la méthode de Monte-Carlo, c'est-à-dire sur un grand nombre de répétitions de la simulation. Or les DFLF que l'on a tracé pour le moment dans ce chapitre ne reposent sur rien de tel. Pour tenter une comparaison entre les DFLF issus de la modélisation d'une part et ceux issus d'un système réel d'autre part, nous allons construire un DFLF moyen.

4.4.1 Diagramme fondamental de ligne ferroviaire moyen : cas réel

Sur la figure [4.7](#) nous avons représenté les huit DFLF moyens pour chaque ligne, chaque direction et pour les cas de figures planifiés et réalisés. L'écart type est également représenté. L'ensemble des données disponibles a été mis à contribution de la manière suivante :

Pour chaque jour nous avons classé les données des DFLF obtenus précédemment par concentration croissante. Nous avons ensuite tracé les points du DFLF en divisant cette concentration en 30 parts égales et en calculant le débit moyen de l'ensemble des points par part de concentration. On considère que les points d'une même journée sont statistiquement liés et on ne calcule pas d'écart type sur une journée. Nous réunissons ensuite les diagrammes de toutes les journées disponibles pour construire le DFLF moyen. On calcule aussi son écart-type entre les différents jours. Nous considérons en effet que les points issus de deux journées différentes ne sont pas statistiquement liés.

Ce DFLF est donc l'analogue des DFLF tracés à partir des modèles présentés dans les chapitres 2 et 3 et construits à partir d'un grand nombre de simulations. Malgré l'accumulation des données d'une trentaine de jours environ pour chaque courbe, les concentrations maximales observées sur le DFLF réalisé de la ligne B restent très rarement atteintes. Il faut donc considérer que les points du diagramme sont moins significatifs pour la ligne B au-dessus d'une concentration qui vaut environ $0,4 \text{ train.km}^{-1}$ car les points tracés au-delà le sont à partir de moins de 8 jours. Sur la ligne B la concentration des trains peut dépasser celle prévu par le plan de transport. Pour la ligne C la courbe reste significative jusqu'au bout, la concentration maximale ne semble pas beaucoup excéder celle prévu par le plan de transport. L'origine de la différence entre les lignes B et C n'a pas été identifiée.

4.4.2 Mise en perspective avec la modélisation

Les DFLF moyennés que l'on veut comparer sont des outils qui ne permettent pas d'observer la dynamique du système comme le permettent les DFLF à la journée où les courbes de la probabilité d'apparition de la congestion au cours du temps. Nous allons faire ici une synthèse des résultats obtenus dans les chapitres 2 et 3 pour ce DFLF.

On a étudié dans ces deux chapitres, les variations de quelques paramètres : la marge de vitesse, le nombre de trains présents sur le réseau et l'amplitude des aléas. On va laisser de côté le rôle de la marge de vitesse pour laquelle on observait des effets bien plus faibles et se concentrer sur les deux autres paramètres.

Concernant les aléas, nous avons également à l'issue de chapitre 3, pu discuter de l'influence d'un aléa décentré sur le trafic en général et sur le DFLF en particulier. Il faut reconnaître que cet effet n'a cependant pas été étudié avec la plus grande rigueur. En effet, avec la force aléatoire, nous n'avons testé qu'un aléa centré, alors qu'avec les temps d'arrêts en gare nous n'avons testé qu'un aléa décentré. Nous n'avons donc pas à proprement parlé testé les effets de la dé-symétrisation du hasard toutes choses égales par ailleurs. Cependant, le raisonnement exposé en 3.3.3 pallie ce manque et permet de comprendre les effets que l'on peut attribuer à cette dé-symétrisation. On a pour l'occasion, proposé de séparer un aléa non centré comme étant la combinaison d'une composante déterministe et d'un aléa centré. Pour comparer les DFLF issus des données avec le modèle des chapitres précédents, on a reproduit ci-après la figure 3.10 (figure 4.8).

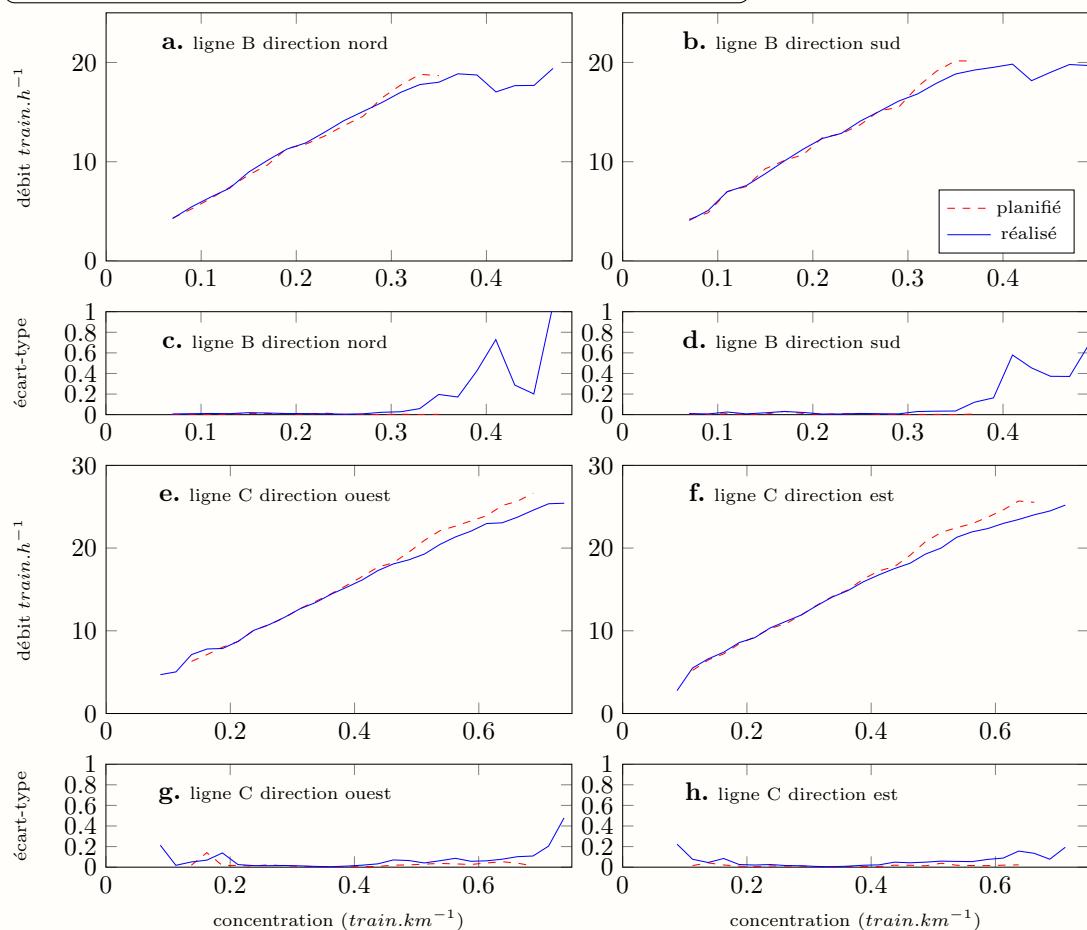
Il semblerait en regardant la figure 4.7 que les effets visibles correspondent avant tout à un effondrement du débit pour les hautes concentrations. La comparaison avec la figure 4.8 nous offre un argument pour soutenir l'hypothèse que l'on observe un phénomène de congestion dû à des aléas. On ne peut pas écarter le caractère non centré de ces aléas, on peut seulement dire que si jamais ces aléas sont non centrés, alors le dé-centrage est insuffisant pour avoir des effets visibles sur le DFLF.

Il faut toutefois rester prudent car le modèle des chapitres 2 et 3 diffère encore de la situation réelle par un certain nombre de caractéristiques que l'on va présenter ici :

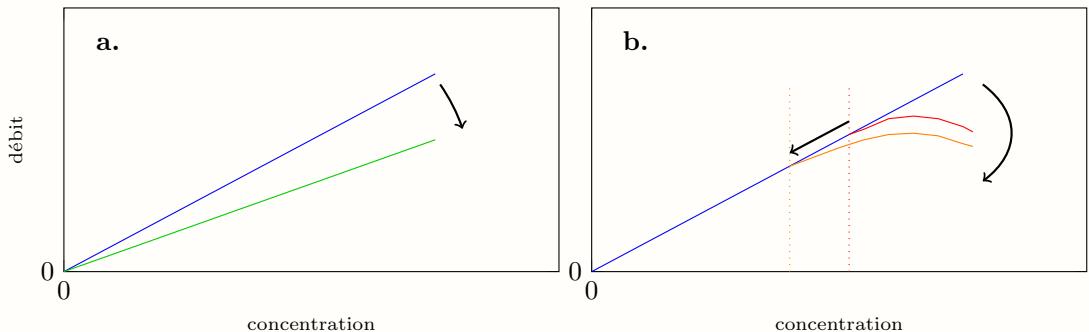
- La portion de la ligne B étudiée n'est pas linéaire mais à topologie en «Y». En revanche, la ligne C est bien linéaire.

- Le modèle des chapitres 2 et 3 est statique du point de vue du nombre de train (ou de la concentration), on ne regarde pas des situations de montée en charge du trafic comme c'est le cas chaque jour sur les lignes étudiées.
- La régulation du trafic n'est pas prise en compte dans le modèle alors qu'elle est présente dans la réalité et qu'elle permet d'empêcher, de limiter ou de résorber des situations de congestion souvent en supprimant des trains initialement prévus par le plan de transport ou en modifiant ce dernier. On a l'intuition que sans cette régulation les DFLF réalisés de la figure 4.7 exhiberaient des chutes de débit plus importantes.

Figure 4.7 - DFLF moyennés sur l'ensemble des jours



DFLF planifiés et réalisés moyennés sur l'ensemble des jours disponibles. Les sous-figures **a** et **b** représentent les DFLF de la ligne B dans les deux directions, tandis que les figures **c** et **d** représentent leurs écart-types. Parmi les jours collectés, trop peu voient leurs concentrations prendre des valeurs supérieures à $0,4 \text{ train.km}^{-1}$ par conséquent la courbe réalisée obtenue cesse d'être significative après ce seuil. Le manque de données après ce seuil explique également les fluctuations de l'écart-type. Les sous-figures **e** et **f** représentent les DFLF de la ligne C dans les deux directions, tandis que les figures **g** et **h** représentent leurs écart-types. Pour la ligne C, la courbe reste significative jusqu'à ses dernières valeurs.

Figure 4.8 - Effets des aléas sur le DFLF


Effets des aléas sur le DFLF. La sous-figure **a** représente les effets sur le DFLF d'une perturbation déterministe : un changement de la pente du diagramme. La sous-figure **b** représente les effets sur le DFLF de la composante aléatoire centrée d'un aléa : un effondrement du débit pour les hautes concentrations.

- Les interactions entre deux trains, que l'on a modélisées, prennent uniquement en compte la signalisation. Or, il existe une autre interaction qui joue un rôle sur les temps d'arrêts. Elle fait intervenir les voyageurs de la manière suivante : si l'écart entre deux trains se creuse, les voyageurs sont plus susceptibles de s'accumuler sur le quai et donc de rallonger le temps d'arrêt en gare du deuxième de ces trains. On a bien modélisé une gestion stochastique de ces temps d'arrêts mais sans leur donner une dépendance à l'écartement horaire entre les trains. Or cette interaction existe.
- Enfin, le plan de transport et l'infrastructure de la ligne que l'on a modélisée ont été idéalisés et ne correspondent pas à la réalité qui n'est pas aussi homogène.

4.5 Discussion et conclusion

Nous avons montré dans ce chapitre qu'un DF local n'est pas un outil qui a du sens pour le ferroviaire en milieu urbain. En revanche, nous avons pu construire un DF non local pour le ferroviaire. Ce genre de diagramme existe depuis un certain temps pour les réseaux routiers, qu'ils soient urbains ou autoroutiers, une riche littérature l'illustre. Ici, nous avons proposé une façon de construire un tel outil à partir de la détection par les balises du passage des trains, le DFLF et nous l'avons appliqué à deux lignes où la fréquence des trains est particulièrement élevée. Avec ce schéma, nous mettons en évidence que la vitesse moyenne des trains est inférieure à celle prévue par le plan de transport.

On a, dans un premier temps, séparé visuellement les jours pour lesquels les DFLF réalisés présentent une forme linéaire, de ceux qui présentent une forme différente. On mettait ainsi en valeur l'existence sur ces derniers d'un fort phénomène de congestion impactant le débit de la ligne.

Nous avons également vérifié que pour les jours classés dans la catégorie des DFLF planifiés linéaires, la régression linéaire proportionnelle (linéaire passant par zéro) était

meilleure que la régression parabolique pour les DFLF théoriques comme pour ceux réalisés. Ces résultats sont rassurants car ils sont cohérents avec la pratique d'une construction du plan de transport à voie libre.

Le concept du DFLF en est néanmoins à ses balbutiements et nécessite un travail d'approfondissement. La recherche présentée dans cette thèse étant seulement la première étape d'un projet plus vaste.

Conclusion

L'objectif de ces travaux de thèse était double. En partant de la constatation de l'existence d'un état du trafic dégradé en routier comme en ferroviaire, nous voulions étudier la faisabilité et les conditions nécessaires à la transposition des outils du trafic routier vers le trafic ferroviaire. L'existence d'un état du trafic dégradé en ferroviaire nous a également amené à nous demander si la modélisation stochastique de perturbations du trafic pouvait, dans un modèle simplifié, permettre de reproduire cet état dégradé du trafic.

Résultats

Nous avons présenté au chapitre 1 une comparaison entre les modes de transport routier et ferroviaire qui met en valeur les éléments à prendre en considération en vue d'une adaptation des outils du trafic routier vers le trafic ferroviaire. En particulier nous avons identifié comme étant problématique les éléments suivants :

- La différence de densité des véhicules ;
- La gestion planifiée du ferroviaire complété par des opérateurs de régulation du trafic qui peuvent modifier le plan de transport en temps réel.
- La nature discrétisée de l'interaction entre véhicules en ferroviaire.

Dans le chapitre 2, nous avons développé un modèle simplifié d'évolution du trafic sur une infrastructure homogène et sans gare. L'originalité de ce modèle réside dans une transcription des règles de conduite des trains sous forme de 3 forces auxquelles un train est soumis. Une de ces forces modélise l'adhérence à un horaire, une autre l'interaction entre les trains et la troisième moins importante le respect des vitesses maximum autorisées. On a montré que l'utilisation d'une quatrième force, aléatoire, qui sert à modéliser les perturbations, permet de reproduire l'apparition d'un état de trafic dégradé. On a défini la congestion comme étant un état du trafic dont la dégradation passe par l'interaction entre les trains. Ce modèle permet d'étudier l'influence de 3 paramètres sur l'apparition de cet état : on a observé que l'augmentation de la concentration des trains ou de l'amplitude des aléas accélère conjointement l'apparition de la congestion et sa gravité. On a également observé que le rôle du troisième paramètre, la marge de vitesse, est secondaire par rapport aux deux premiers. La dynamique de l'apparition de la congestion a été étudiée dans le cadre de la théorie des processus de Markov. La quantification de la congestion nous a permis de définir deux sous-états de la congestion :

- La congestion résiliente, où les retards, causés par les aléas et amplifiés par les interactions entre trains, sont résorbés grâce à la marge de vitesse. La vitesse moyenne des trains en congestion résiliente est la même que sans congestion.
- La congestion dégradée, où les retards ne peuvent être rattrapés et où la vitesse moyenne des trains est plus faible que sans la congestion. Dans cette configuration et sans intervention extérieure, les retards s'accumulent.

Finalement, le tracé du diagramme fondamental de ligne ferroviaire nous a permis de visualiser la forme que prenait la courbe du débit moyen en fonction de la concentration moyenne et de l'amplitude des aléas.

Dans le chapitre 3, nous avons modifié le modèle précédent afin tenir compte de la présence des gares. Nous n'avons étudié que les effets de deux des trois variables précédentes, le nombre de trains et l'amplitude des aléas. En contrepartie, cette variante du modèle initial était l'occasion d'introduire un nouveau type d'aléa sur les temps de stationnement des trains et de comparer les effets de ce dernier avec l'aléa modélisé par une force. Cette comparaison a permis d'observer une dynamique plus lente d'apparition de la congestion dans le cas où l'aléa est modélisé par des temps d'arrêt en gare. Un raisonnement analytique nous a permis de comprendre le rôle de l'hétérogénéité de la ligne sur l'apparition de la congestion et en particulier sur sa dynamique. L'étude des vitesses moyennes de la ligne était l'occasion d'observer les deux mêmes sous-états de la congestion que dans le chapitre précédent. Le tracé des diagrammes fondamentaux de ligne ferroviaire a permis de mettre en évidence, dans le cas d'aléas modélisés par des variations des temps de stationnement, une baisse de la vitesse moyenne sans interaction entre les trains. Cette dégradation du trafic, qui ne correspond pas à la définition de la congestion, est attribuée à la nature décentrée des aléas utilisés.

Le chapitre 4 était l'occasion de comparer les résultats obtenus de manière théorique et simulés dans les deux chapitres précédents grâce à une analyse des données de trafic ferroviaire des lignes B et C du RER, le réseau ferré d'Ile-de-France. Après avoir présenté ce jeu de données, nous avons mis en évidence l'impossibilité de tracer un DF local à partir des données disponibles. Nous avons cependant montré l'inutilité d'un tel diagramme. En nous inspirant du DFR en trafic routier, nous avons montré qu'il était possible de construire un diagramme fondamental non local en ferroviaire qui agrège les données disponibles sur une portion de ligne. La première version de ce diagramme, le diagramme fondamental de ligne ferroviaire, est tracé pour chaque jour disponible. Il a permis de mettre en évidence, expérimentalement, la proportionnalité entre le débit et la concentration prévue par le plan de transport, mais aussi l'existence de périodes de dégradation du trafic. L'observation des formes diverses et non répétées des diagrammes fondamentaux de ligne ferroviaire correspondant à ces dégradations du trafic est un argument en faveur d'une modélisation stochastique des aléas introduite au chapitre 2. En regroupant les données de tous les jours disponibles en un seul diagramme fondamental de ligne ferroviaire, on a ouvert la voie à la comparaison avec les diagrammes fondamentaux de ligne ferroviaire simulés sur un grand nombre de répétitions d'une simulation des modèles des deux

chapitres précédents. Cette comparaison qualitative renforce l'hypothèse selon laquelle la modélisation stochastique des perturbations permet de reproduire la dégradation de l'état du trafic observé avec les données.

Perspectives

La nature exploratoire du travail de thèse réalisé permet d'ouvrir, grâce aux résultats obtenus, de nombreuses pistes de réflexion pour la modernisation des méthodes d'exploitation du ferroviaire en zone dense. Cependant la consolidation des résultats obtenus nécessite encore beaucoup d'efforts. En particulier, parmi les points susceptibles d'être améliorés et les pistes de recherche prometteuses, on distingue :

- La calibration du modèle sur un cas d'étude n'a pas été réalisée et la comparaison entre le modèle et les données n'est que qualitative dans l'état actuel des travaux. Une analyse poussée des données précises d'une ligne est nécessaire en vue de cette calibration.
- Si l'effet de l'amplitude des aléas a bien été étudié, les effets des autres caractéristiques des aléas n'ont que peu, voire pas été explorés. En variant la nature des aléas dans le chapitre 3, on a également étudié le dé-centrage de la source d'aléa. En toute rigueur, une étude avec des aléas centrés et décentrés pour chaque type d'aléa testé est nécessaire, d'autres formes de distribution pourraient également être testée. Nous n'avons pas exploré non plus le rôle du temps de corrélation de l'aléa modélisé par une force. Une étude analytique de son rôle est sans doute possible et fournirait une compréhension meilleure qu'une étude basée sur des simulations.
- Le modèle n'a pas permis d'explorer le rôle du nombre de passagers en attente dans l'interaction entre les trains qui n'a été modélisé pour ne tenir compte que de la signalisation. Or l'accumulation des passagers sur les quais dépend de l'écart horaire entre deux trains. Une étude complète du rôle des passagers dans la génération de la congestion serait un progrès indéniable. Une telle étude pourrait s'inspirer d'un phénomène similaire pour les bus.
- Un effort particulier pour reproduire les diagrammes fondamentaux de ligne ferroviaire obtenus sur d'autres lignes doit être fait. La recherche d'un jeu de données plus riche et moins bruité devrait permettre de réduire la dispersion des diagrammes fondamentaux de ligne ferroviaire basés sur des données empiriques. La recherche d'une concentration critique en analogie avec le trafic routier pourrait être mené en travaillant sur un ensemble de données plus grand.
- En ce qui concerne l'intérêt opérationnel de ces travaux, nous pouvons espérer que la transposition de résultats plus récents issus de la théorie de la circulation routière pourrait nous aider à trouver de meilleures façons d'anticiper ou de contrôler la congestion dans un réseau de chemin de fer et en particulier sur le réseau Transilien dans la région de parisienne.
- Enfin, le test de la reproductibilité des résultats obtenus étant une des bases de la méthode scientifique, une validation rigoureuse de ces résultats nécessiterait un travail de reproduction par une équipe de recherche extérieure, et pas seulement une

lecture critique. Ce dernier point, indépendant de notre volonté, devait cependant être souligné dans la mesure où le manque de reproductibilité des résultats est l'un des principaux maux dont souffre la recherche scientifique de nos jours.

Bibliographie

- ALLENBACH, J., P. CHAPAS, M. COMTE et R. KALLER. 2008, *Traction électrique*, vol. 1, Presses polytechniques et universitaires romandes, ISBN 9782880746742. [8](#)
- BANKS, J. H. 1989, «Freeway speed-flow-concentration relationships : more evidence and interpretations (with discussion and closure)», *Transportation Research Record*, , n° 1225. [34](#)
- BECK, J. R. et S. G. PAUKER. 1983, «The markov process in medical prognosis», *Medical Decision Making*, vol. 3, n° 4, p. 419–458. [62](#)
- BENTZ, C. 2014, «Trains transiliens : dans quelle mesure les données de retour d'expérience permettent-elles d'améliorer le service fourni aux voyageurs ?», . [29](#)
- BÜKER, T. et B. SEYBOLD. 2012, «Stochastic modelling of delay propagation in large networks», *Journal of Rail Transport Planning & Management*, vol. 2, n° 1, p. 34–50. [46](#), [62](#)
- CACCHIANI, V., D. HUISMAN, M. KIDD, L. KROON, P. TOTH, L. VEELENTURF et J. WAGENAAR. 2014, «An overview of recovery models and algorithms for real-time railway rescheduling», *Transportation Research Part B : Methodological*, vol. 63, p. 15 – 37. [29](#)
- CAREY, M. et A. KWIECIŃSKI. 1994, «Stochastic approximation to the effects of headways on knock-on delays of trains», *Transportation Research Part B : Methodological*, vol. 28, n° 4, p. 251–267. [46](#)
- CAREY, M. et A. KWIECIŃSKI. 1995, «Properties of expected costs and performance measures in stochastic models of scheduled transport», *European Journal of Operational Research*, vol. 83, n° 1, p. 182–199. [46](#)
- CHIABAUT, N. 2009, «Écoulement du trafic routier et états d'équilibre en congestion», . [31](#), [34](#), [35](#)
- CUNIASSE, P.-A., C. BUISSON, J. RODRIGUEZ, E. TEOUL et D. DE ALMEIDA. 2015, «Analyzing railroad congestion in a dense urban network through the use of a road traffic network fundamental diagram concept», *Public Transport*, p. 1–13. [113](#)
- DAGANZO, C. 2007, «Urban gridlock : Macroscopic modeling and mitigation approaches», *Transportation Research B*, , n° 41(1), 49-62. [114](#)
- DAGANZO, C. F. 1997, «A continuum theory of traffic dynamics for freeways with special lanes», *Transportation Research Part B : Methodological*, vol. 31, n° 2, p. 83–102. [31](#)
- DAGANZO, C. F. et N. GEROLIMINIS. 2008, «An analytical approximation for the macroscopic fundamental diagram of urban traffic», *Transportation Research Part B : Methodological*, vol. 42, n° 9, p. 771 – 781, ISSN 0191-2615. [114](#)
- DIU, B., C. GUTHMANN, D. LEDERER et B. ROULET. 1989a, *Physique statistique, Chapitre 1*, Hermann, Paris, p. 3–8. [32](#), [33](#)

- DIU, B., C. GUTHMANN, D. LEDERER et B. ROULET. 1989b, *Physique statistique, Chapitre 4*, Hermann, Paris, p. 562–571. [32](#)
- EDIE, L. C. 1963, «Discussion of traffic stream measurements and definitions», *ISTTT*, p. 139–154. [31](#)
- ENGELHARDT-FUNKE, O. et M. KOLONKO. 2002, «Simulating delays for realistic time-table optimization», dans *Operations Research Proceedings 2001*, Springer, p. 9–15. [45](#)
- GÉLY, L., G. DESSAGNE et C. LÉRIN. 2006, «Modelling train re-scheduling with optimization and operational research techniques : results and applications at sncf», dans *World Congress on Railway Research*. [29](#)
- GEROLIMINIS, N. et C. DAGANZO. 2007, «Macroscopic modeling of traffic in cities», *Transportation Research Board*, , n° 07-0413. [114](#)
- GEROLIMINIS, N. et C. F. DAGANZO. 2008, «Existence of urban-scale macroscopic fundamental diagrams : Some experimental findings», *Transportation Research Part B : Methodological*, vol. 42, n° 9, p. 759 – 770, ISSN 0191-2615. [114](#)
- GEROLIMINIS, N. et J. SUN. 2011, «Properties of a well-defined macroscopic fundamental diagram for urban traffic», *Transportation Research Part B : Methodological*, vol. 45, n° 3, p. 605 – 617, ISSN 0191-2615. [114](#)
- GODFREY, J. 1969, «The mechanism of a road network», *Traffic Engineering and Control*, , n° Vol. 11, No. 7, pp. 323-327. [113](#)
- GREENSHIELDS, B., W. CHANNING et H. MILLER. 1935, «A study of traffic capacity», dans *Highway research board proceedings*, vol. 1935, National Research Council (USA), Highway Research Board. [34](#)
- HELBING, D. 2001, «Traffic and related self-driven many-particle systems», *Rev. Mod. Phys.*, vol. 73, doi :10.1103/RevModPhys.73.1067, p. 1067–1141. [39](#)
- HERMAN, R. et I. PRIGOGINE. 1979, «A two-fluid approach to town traffic», *Science*, , n° 204, pp. 148-151. [113](#)
- HIGGINS, A. et E. KOZAN. 1998, «Modeling train delays in urban networks», *Transportation Science*, vol. 32, n° 4, p. 346–357. [46](#)
- HIGGINS, A., E. KOZAN et L. FERREIRA. 1996, «Optimal scheduling of trains on a single line track», *Transportation Research Part B : Methodological*, vol. 30, n° 2, p. 147 – 161. [29](#)
- HOLLNAGEL, E., B. JOURNÉ et H. LAROCHE. 2009, «La fiabilité et la résilience comme dimensions de la performance organisationnelle», *Management*, vol. 12, n° 4, p. 224–229. [14](#)
- LAI, Y.-C. R., Y.-H. LIU et T.-Y. LIN. 2012, «Development of base train equivalents to standardize trains for capacity analysis», *Transportation Research Record : Journal of the Transportation Research Board*, vol. 2289, n° 1, p. 119–125. [47](#)

- LAMALLE, U. 1947, *Cours d'exploitation des chemins de fer*, n° vol. 1 à 2 dans Cours d'exploitation des chemins de fer, Librairie Universitaire. [8](#)
- LECLERC, L. 2009, «Le modèle lwr : théorie, confrontation expérimentale et applications au milieu urbain», . [32](#)
- LIEBCHEN, C., M. SCHACHTEBECK, A. SCHÖBEL, S. STILLER et A. PRIGGE. 2010, «Computing delay resistant railway timetables», *Computers and Operations Research*, vol. 37, n° 5, p. 857 – 868. Disruption Management. [29](#)
- LIGHTHILL, M. J. et G. B. WHITHAM. 1955, «On kinematic waves. ii. a theory of traffic flow on long crowded roads», dans *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, vol. 229, The Royal Society, p. 317–345. [30, 47](#)
- MAHMASSANI, H., J. C. WILLIAMS et R. HERMAN. 1984, «Investigations of network-level traffic flow relationships : Some simulation results», *Transportation Research Record* 971, , n° pp. 121-130. [114](#)
- MEESTER, L. E. et S. MUNS. 2007, «Stochastic delay propagation in railway networks and phase-type distributions», *Transportation Research Part B : Methodological*, vol. 41, n° 2, p. 218–230. [46, 62](#)
- MÜLHANS, E. et C.-R. WARNINGHOFF. 1990, «Neue strassenhanen ins umland ?», *Internationales Verkehrswesen*, vol. 42, n° 5, p. 286–292. [45](#)
- NEWELL, G. F. 1993, «A simplified theory of kinematic waves in highway traffic, part i : General theory», *Transportation Research Part B : Methodological*, vol. 27, n° 4, p. 281–287. [34](#)
- POINCARÉ, H. 1908, «Science et méthode, livre premier», . [44](#)
- RICHARDS, P. I. 1956, «Shock waves on the highway», *Operations research*, vol. 4, n° 1, p. 42–51. [30, 47](#)
- SABERI, M. et H. S. MAHMASSANI. 2012, «Exploring the properties of network-wide flow-density relations in a freeway network», *Transportation Research Record*, , n° No. 2315, pp. 153-163. [114](#)
- SABERI, M., H. S. MAHMASSANI, T. HOU et A. ZOCKAIE. 2014, «Estimating network fundamental diagram using three-dimensional vehicle trajectories : Extending edie's definitions of traffic flow variables to networks», *Transportation Research Record : Journal of the Transportation Research Board*, vol. 2422, p. 12–20. [114](#)
- SCHWANHÄUSSER, W. 1974, *Die Bemessung der Pufferzeiten im Fahrplangefüge der Eisenbahn*, thèse de doctorat, Verkehrswissenschaftliches Institut der Rheinisch-Westfälischen Technischen Hochschule. [45](#)
- SZÁSZ, D. 1996, «Boltzmann's ergodic hypothesis, a conjecture for centuries ?», *Studia Scientiarum Mathematicarum Hungarica*, vol. 31, n° 1, p. 299–322. [32](#)

- TÖYLI, H. L., J. LAURI OJALA, A. WIELAND et C. MARCUS WALLENBURG. 2013, «The influence of relational competencies on supply chain resilience : a relational view», *International Journal of Physical Distribution & Logistics Management*, vol. 43, n° 4, p. 300–320. [14](#)
- VICSEK, T. et A. ZAFEIRIS. 2012, «Collective motion», *Physics Reports*, vol. 517, n° 3, p. 71–140. [39](#)
- WEIGAND, W. 1981, «Verspätungsübertragungen in fernverkehrsnetzen», *Eisenbahntechnische Rundschau (ETR)*, vol. 30, p. 915–920. [45](#)
- YUAN, J. 2006, *Stochastic modelling of train delays and delay propagation in stations*, vol. 2006, Eburon Uitgeverij BV. [45](#)
- YUAN, J., R. GOVERDE et I. HANSEN. 2002, «Propagation of train delays in stations», *Pубlication of: WIT Press*. [46](#)

Annexe A

Détermination des paramètres de la force horaire, F_{horaire}

Dans la partie 2.2.3, on cherche à déterminer les valeurs de a_{elastiq} , a_{friction} , h_d et de t_{friction} . On va voir que l'on est obligé de fixer h_d et t_{car} nous même.

On part des 5 points suivants :

- (A.1) La définition du délai Δh : $\Delta h = \Delta h(0) + \int_0^t (\nu(x) - \nu_{\text{plan}}) / \nu_{\text{plan}} \cdot dx$.
- (A.2) La seconde loi de Newton : $m \cdot a = F_h$.
- (A.3) L'hypothèse de petites variations de Δh et de $\partial_t \Delta h$ et donc la linéarisation de la fonction \tanh .
- (A.4) La relation : $a_{\text{elastiq}} + a_{\text{friction}} = a_{\text{max}}$. C'est une condition suffisante pour que la force horaire ne dépasse pas la valeur de l'accélération maximum a_{max} . De plus on a nécessairement $a_{\text{elastiq}}, a_{\text{friction}} > 0$.
- (A.5) On a avec la linéarisation de la \tanh une vraie réponse proportionnelle dérivée. On choisit de prendre les paramètres qui permettent que cette réponse soit la plus rapide possible sans oscillations. C'est la solution critique de l'équation 2 en réponse au délai. Ce choix se fonde sur l'absence observée d'oscillations marquées du délai lorsqu'un conducteur corrige un retard ou une avance.

On a (A.2) : $m \cdot a = -m \cdot a_{\text{elastiq}} \cdot \tanh(\Delta h / h_d) - m \cdot a_{\text{friction}} \cdot \tanh(t_{\text{friction}} \cdot \partial_t \Delta h / h_d)$

On linéarise (A.3) :

$$\begin{aligned} \Rightarrow m \cdot a &= -m \cdot a_{\text{elastiq}} \cdot \Delta h / h_d - m \cdot a_{\text{friction}} \cdot t_{\text{friction}} \cdot \partial_t \Delta h / h_d \\ \Leftrightarrow a &= -a_{\text{elastiq}} \cdot \Delta h / h_d - a_{\text{friction}} \cdot t_{\text{friction}} \cdot \partial_t \Delta h / h_d \end{aligned}$$

On a (A.1) : $\Delta h = \Delta h(0) + \int_0^t (\nu(x) - \nu_{\text{plan}}) / \nu_{\text{plan}} \cdot dx$

(A.1) et (A.2) : $a = -\frac{a_{\text{elastiq}}}{h_d} \cdot \left(\Delta h(0) + \int_0^t \frac{\nu(x) - \nu_{\text{plan}}}{\nu_{\text{plan}}} \cdot dx \right) - \frac{a_{\text{friction}} \cdot t_{\text{friction}}}{h_d} \cdot \frac{\nu(x) - \nu_{\text{plan}}}{\nu_{\text{plan}}}$

En posant : $\alpha_{\text{elastiq}} = a_{\text{elastiq}} / h_d$ et $\alpha_{\text{friction}} = a_{\text{friction}} \cdot t_{\text{friction}} / h_d$

$$\begin{aligned}
 & \Leftrightarrow a = -\alpha_{elastiq} \cdot \left(\Delta h(0) + \int_0^t \frac{v(x) - v_{plan}}{v_{plan}} \cdot dx \right) - \alpha_{friction} \cdot \frac{v(x) - v_{plan}}{v_{plan}} \\
 & \Leftrightarrow \partial_t v + \alpha_{elastiq} \cdot \left(\Delta h(0) + \int_0^t \frac{v(x) - v_{plan}}{v_{plan}} \cdot dx \right) + \alpha_{friction} \cdot \frac{v(x) - v_{plan}}{v_{plan}} = 0 \\
 & \Leftrightarrow \partial_t v + \frac{\alpha_{friction}}{v_{plan}} \cdot v + \frac{\alpha_{elastiq}}{v_{plan}} \cdot \int_0^t v(x) \cdot dx = -\alpha_{elastiq} (\Delta h(0) - t) + \alpha_{friction}
 \end{aligned}$$

On dérive :

$$\begin{aligned}
 & \Rightarrow \partial_t^2 v + \frac{\alpha_{friction}}{v_{plan}} \cdot \partial_t v + \frac{\alpha_{elastiq}}{v_{plan}} \cdot v = \alpha_{elastiq} \cdot \frac{v_{plan}}{v_{plan}} \\
 & \text{(A.5)} \Leftrightarrow \Delta = 0 = \frac{\alpha_{friction}^2}{v_{plan}^2} - 4 \cdot \frac{\alpha_{elastiq}}{v_{plan}} \quad ; \quad t_{car} = \frac{2 \cdot v_{plan}}{\alpha_{friction}}
 \end{aligned}$$

Avec t_{car} (qui est négatif) le temps caractéristique de convergence vers la vitesse v_{plan} . Il est important de noter que la vitesse planifiée n'est pas forcément une constante dans le cas général. Il est de plus sage de prendre t_{car} indépendant de v_{plan} . Il est évident que ce temps pourra en réalité être plus long si v_{plan} est élevé ou si la vitesse initiale est trop éloignée de v_{plan} mais c'est le cas grâce à la non-linéarité de la fonction $\tanh(x)$. Avant de choisir la valeur de t_{car} qui va nous donner $\alpha_{elastiq}$ et $\alpha_{friction}$, revenons justement au système non linéarisé pour déterminer les valeurs de h_d et de $t_{friction}$.

h_d représente le retard ou l'avance qu'un conducteur commencerait à juger assez importante pour réagir au maximum des capacités de sa machine (saturation de la fonction). Au-delà d'un tel retard (ou avance) t_{car} augmente. $h_d / t_{friction}$ représente le taux de rattrapage (en valeur absolue) à partir duquel on sature. Cette saturation va également impacter t_{car} .

On a (A.4) : $a_{elastiq} + a_{friction} = a_{max}$; $a_{elastiq}, a_{friction} > 0$

$$a_{elastiq} = \alpha_{elastiq} \cdot h_d \quad ; \quad a_{friction} = \alpha_{friction} \cdot h_d / t_{friction}$$

avec (A.5) :

$$\alpha_{elastiq} = \alpha_{friction}^2 / 4 \cdot v_{plan} = v_{plan} / t_{car}^2$$

$$\alpha_{friction} = -2v_{plan} / t_{car}$$

$t_{friction}$ est fixé par h_d :

$$\begin{aligned}
 & \alpha_{friction} = a_{friction} \cdot t_{friction} / h_d \\
 & \Leftrightarrow t_{friction} = \frac{\alpha_{friction} \cdot h_d}{a_{friction}} \\
 & t_{friction} = \frac{\alpha_{friction} \cdot h_d}{a_{max} - \alpha_{elastiq} \cdot h_d} = 2 \frac{v_{plan} \cdot h_d}{a_{max} \cdot t_{car} - v_{plan} \cdot h_d / t_{car}}
 \end{aligned}$$

$$\begin{aligned}
 (\text{A.4}) \Rightarrow a_{friction} &= a_{max} - \alpha_{elastiq} \cdot h_d \quad ; \quad 0 < a_{friction} < a_{max} \\
 \Rightarrow \alpha_{elastiq} \cdot h_d &< a_{max} \\
 \Rightarrow h_d &< \frac{a_{max} \cdot t_{car}^2}{v_{plan}}
 \end{aligned}$$

Application numérique :

On a : $v_{plan} = 85 \text{ km.h}^{-1} \approx 23,6 \text{ m.s}^{-1}$.

On choisit : $h_d = 180 \text{ s}$, $a_{max} = 0,6 \text{ m.s}^{-2}$ et $t_{car} = -89,35 \text{ s}$ ¹

On trouve : $t_{friction} = 1255 \text{ s}$, $a_{elastiq} = 0,5247 \text{ m.s}^{-2}$ et $a_{friction} = 0,0758 \text{ m.s}^{-2}$

1. h_d est choisi comme expliqué en 2.2.3. t_{car} n'a pas une valeur entière car dans un premier temps nous avons fixés $t_{friction}$ pour que $t_{car} \approx -90 \text{ s}$.

Annexe B

Détermination des paramètres de la force aléatoire

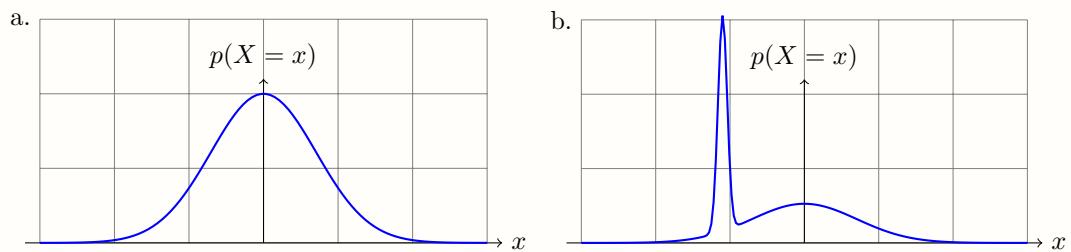
Dans la partie 2.2.3, on cherche à construire une force aléatoire F_{rand} . Cette force est calculée aléatoirement à chaque pas de temps, il convient d'imposer à cette force un temps d'auto-corrélation t_{corr} nettement supérieur au pas de temps. t_{corr} étant environ le temps qu'il faut pour que la chance que le résultat du générateur aléatoire change soit de 50%. Pour ce faire on utilise la méthode suivante :

Le premier tirage est fait selon une loi de distribution $d_0(x)$ dont les réalisations sont incluses dans l'intervalle $[a_{min}, a_{max}]$. On construit la dépendance du tirage i au tirage $i - 1$ en combinant linéairement $d_0(x)$ et le Dirac centré sur la réalisation du tirage $i - 1$. Les deux distributions correspondantes sont présentées en figure B.1.

$$d_i(x) = (1 - p) \cdot d_0(x) + p \cdot \delta_{r(i-1)} \quad (\text{B.1})$$

Le choix de p permet de régler le temps de corrélation, il dépend du pas de temps Dt . On a : $p = e^{\ln(1/2) \cdot Dt / t_{corr}}$

Figure B.1 - distribution de F_{rand}



a. : exemple de loi de probabilité initiale. b. : loi probabilité au tirage $n + 1$ le pic correspond à la valeur du tirage précédent, le rapport de l'aire du pic sur l'aire totale permet de régler le temps de corrélation.

Application numérique :

Si on veut $t_{corr} = 300\text{s}$ avec $Dt = 0.1\text{s}$ il faut $p = 0.999769$

Contrairement à la démonstration qui se fait dans le cas général, dans cette thèse, d_0 sera une distribution uniforme. Pour ne pas qu'on impute le retard et la congestion à une espérance négative de d_0 , cette distribution est centrée en 0 bien que cela puisse paraître optimiste.

Annexe C

Mise en équation de la compétition entre la force aléatoire et la force horaire

Comme dans le chapitre 2, on se place dans le cadre où F_{rand} garde la même valeur, que l'on appelle f , pendant un certain temps. f vérifie la relation :

$$-m \cdot a_{r\ max} \leq f \leq m \cdot a_{r\ max} \quad (\text{C.1})$$

La seconde loi de Newton appliquée à un train s'écrit alors :

$$\begin{aligned} m \cdot a(t) &= F_{rand} + F_{horaire} \\ \Leftrightarrow m \cdot a(t) &= f + F_{elastiq} + F_{friction} \\ \Leftrightarrow m \cdot a(t) &= f - m \cdot a_{elastiq} \cdot \tanh\left(\frac{\Delta h}{h_d}\right) - m \cdot a_{friction} \cdot \tanh\left(\frac{\partial_t(\Delta h) \cdot t_{friction}}{h_d}\right) \end{aligned} \quad (\text{C.2})$$

On reprend la définition du délai :

$$\Delta h = \Delta h(0) + \int_0^t (v(x) - v_{plan}) / v_{plan} \cdot dx \quad (\text{C.3})$$

On dérive deux fois la relation (C.3) :

$$\begin{aligned} \partial_t(\Delta h) &= \frac{v(t) - v_{plan}}{v_{plan}} \\ \partial_t^2(\Delta h) &= \frac{a(t)}{v_{plan}} \end{aligned}$$

On injecte dans la relation (C.2) :

$$\begin{aligned} v_{plan} \cdot \partial_t^2(\Delta h) &= \frac{f}{m} - a_{elastiq} \cdot \tanh\left(\frac{\Delta h}{h_d}\right) - a_{friction} \cdot \tanh\left(\frac{\partial_t(\Delta h) \cdot t_{friction}}{h_d}\right) \\ \Leftrightarrow \partial_t^2(\Delta h) + \frac{a_{friction}}{v_{plan}} \cdot \tanh\left(\frac{\partial_t(\Delta h) \cdot t_{friction}}{h_d}\right) &+ \frac{a_{elastiq}}{v_{plan}} \cdot \tanh\left(\frac{\Delta h}{h_d}\right) - \frac{f}{m \cdot v_{plan}} = 0 \end{aligned} \quad (\text{C.4})$$

On obtient alors une équation différentielle non linéaire du second ordre dont la variable est le délai Δh . Considérons que l'état d'équilibre est atteint ; on a :

$$\begin{aligned} \text{(équilibre)} &\Leftrightarrow \partial_t^2(\Delta h) = 0 \text{ et } \partial_t(\Delta h) = 0 \\ \text{(équilibre)} + (\text{C.4}) &\Rightarrow \frac{a_{elastiq}}{v_{plan}} \cdot \tanh\left(\frac{\Delta h}{h_d}\right) - \frac{f}{m \cdot v_{plan}} = 0 \end{aligned}$$

La valeur que prend le délai à l'équilibre Δh_{equ} s'écrit :

$$\Delta h_{equ} = h_d \cdot \operatorname{arctanh}\left(\frac{f}{m \cdot a_{elastiq}}\right)$$

À l'équilibre, on peut donc considérer que :

$$(\text{C.1}) \Leftrightarrow -h_d \cdot \operatorname{arctanh}\left(\frac{-m \cdot a_{r\ max}}{m \cdot a_{elastiq}}\right) \leq \Delta h_{equ} \leq h_d \cdot \operatorname{arctanh}\left(\frac{m \cdot a_{r\ max}}{m \cdot a_{elastiq}}\right)$$

Selon l'équation (C.4), le risque de dépassement de ces valeurs en dehors de l'équilibre est exclu : selon les valeurs de f , en analogie avec l'équation différentielle linéaire du second ordre correspondante :

$$v_{plan} \cdot \partial_t^2(\Delta h) = \frac{f}{m} - a_{elastiq} \cdot \frac{\Delta h}{h_d} - a_{friction} \cdot \frac{\partial_t(\Delta h) \cdot t_{friction}}{h_d}$$

Il existe deux régimes de convergence vers l'équilibre :

- Le cas $f = 0$ correspond par construction de la force $F_{horaire}$ au régime critique (convergence la plus rapide sans oscillations) (voir annexe A).
- Les cas $f < 0$ et $f > 0$. On ne peut plus linéariser les équations autour de $\Delta h = 0$, il faut linéariser autour du nouvel équilibre : $\Delta h_{equ} = -\operatorname{arctanh}\left(\frac{f}{a_{elastiq}}\right)$. Cela revient à diminuer la valeur du discriminant de l'équation (4) et donc à la rendre positive. Il n'y a donc pas de régimes oscillant possible.

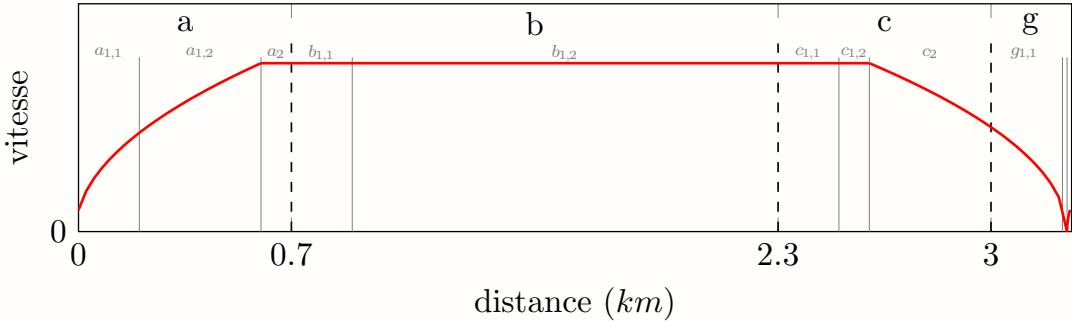
Cependant, en raison de la présence de la force de sécurité $F_{v\ max}$, de la discréétisation numérique et du fait que F_{rand} n'est pas fixée, un petit dépassement des valeurs limites de Δh_{equ} ne peut être a priori totalement exclu.

Annexe D

Calcul d'optimisation de la taille des cantons

En 3.1.1 on cherche à dimensionner les cantons a , b , c , et g afin de maximiser la capacité et de diminuer le nombre de cantons au kilomètre. Pour cela on a posé le système d'équation (3.2), on va ici le résoudre. On a décomposé le motif de cantons $abcg$ en un ensemble de segments de cantons : $a_{1,1}$, $a_{1,2}$, a_2 , $b_{1,1}$, $b_{1,2}$, $c_{1,1}$, $c_{1,2}$, c_2 , $g_{1,1}$, $g_{1,2}$, $stop$ et g_2 . Cette décomposition est visible dans le tableau 3.1. On reproduit sur la figure D.1, la figure 3.2 pour aider le lecteur à suivre les calculs.

Figure D.1 - Position des cantons et marche du train



Ce graphique à l'échelle représente la position des cantons et des segments de cantons d'un motif $abcg$ superposés avec la marche du train. Les segments $g_{1,2}$ et g_2 sont trop petits pour être représentés.

$$(3.2) \Leftrightarrow \begin{cases} t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + t_p(b_{1,1}) + t_p(b_{1,2}) + t_p(c_{1,1}) \\ \quad + t_p(b_{1,1}) + t_p(b_{1,2}) + t_p(c_{1,1}) + t_p(c_{1,2}) + t_p(c_2) + t_p(g_{1,1}) \\ \quad + t_p(c_{1,1}) + t_p(c_{1,2}) + t_p(c_2) + t_p(g_{1,1}) + t_p(g_{1,2}) + t_{stop} + t_p(g_2) = t \\ t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + t_p(b_{1,1}) \\ \quad + t_p(g_{1,1}) + t_p(g_{1,2}) + t_{stop} + t_p(g_2) = t \end{cases} = t$$

Pour résoudre ce système, on commence par exprimer le plus de variables $t_p()$ possible en fonction de variables connues. On connaît :

$$l_{a\ 1,1} = l_{b\ 1,1} = l_{c\ 1,1} = l_{g\ 1,1} = l_T \quad (\text{D.1})$$

Les phases d'accélération et de freinage s'effectuent à : $a = \pm|a_{plan}|$. On intègre deux fois l'accélération :

$$\begin{aligned} a &= \partial_t v = \partial_t^2 x \quad \text{avec } v \text{ la vitesse et } x \text{ la position.} \\ x &= x_0 + v_0 \cdot t + 1/2 \cdot a_{plan} \cdot t^2 \end{aligned}$$

Sur une distance d correspondant :

- Soit à une phase d'accélération dont la vitesse initiale est nulle.
- Soit par symétrie à une phase de freinage dont la vitesse finale est nulle.

On a $x = d$, $x_0 = 0$ et $v_0 = 0$ on trouve donc que :

$$v = a_{plan} \cdot t \quad (\text{D.2})$$

$$d = 1/2 \cdot a_{plan} \cdot t^2 \quad (\text{D.3})$$

$$\text{soit : } t_p(d) = \sqrt{\frac{2 \cdot d}{a_{plan}}} \quad (\text{D.4})$$

On appelle $v_{plan\ max}$ la vitesse atteinte à la fin de l'accélération ou au début du freinage car elle correspond à la vitesse planifiée maximum. On peut donc à partir des équations (D.2) et (D.3) en identifiant v à $v_{plan\ max}$ et d à $(l_{g\ 2} + l_{a\ 1,1} + l_{a\ 1,2})$ puis à $(l_{c\ 2} + l_{g\ 1,1} + l_{g\ 1,2})$, écrire l'expression de $v_{plan\ max}$:

$$(D.2) + (D.3) \Rightarrow v_{plan\ max}^2 = 2 \cdot a_{plan} \cdot (l_{g\ 2} + l_{a\ 1,1} + l_{a\ 1,2}) = 2 \cdot a_{plan} \cdot (l_{c\ 2} + l_{g\ 1,1} + l_{g\ 1,2})$$

Puis en injectant (D.1) :

$$\Rightarrow v_{plan\ max}^2 = 2 \cdot a_{plan} \cdot (l_{g\ 2} + l_T + l_{a\ 1,2}) = 2 \cdot a_{plan} \cdot (l_{c\ 2} + l_T + l_{g\ 1,2})$$

$$\Rightarrow \begin{cases} l_{a\ 1,2} = \frac{v_{plan\ max}^2}{2 \cdot a_{plan}} - l_{g\ 2} - l_T \\ l_{c\ 2} = \frac{v_{plan\ max}^2}{2 \cdot a_{plan}} - l_{g\ 1,2} - l_T \end{cases} \quad (\text{D.5})$$

Les ensembles de segments de cantons suivants correspondent bien aux phases énoncées ci-avant : $(g_2 + a_{1,1} + a_{1,2})$, $(g_2 + a_{1,1})$ et (g_2) d'une part pour l'accélération et $(c_2 + g_{1,1} + g_{1,2})$ et $(g_{1,1} + g_{1,2})$ et $(g_{1,2})$ pour le freinage d'autre part. On peut donc écrire les égalités dans deux systèmes, un pour les accélérations et un pour les freinages :

Pour la phase d'accélération :

$$\begin{cases} t_p(g_2 + a_{1,1} + a_{1,2}) = \sqrt{\frac{2 \cdot (l_g 2 + l_a 1,1 + l_a 1,2)}{a_{plan}}} \\ t_p(g_2 + a_{1,1}) = \sqrt{\frac{2 \cdot (l_g 2 + l_a 1,1)}{a_{plan}}} \\ t_p(g_2) = \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \end{cases}$$

$$\Leftrightarrow \begin{cases} t_p(a_{1,2}) = \sqrt{\frac{2 \cdot (l_g 2 + l_a 1,1 + l_a 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot (l_g 2 + l_a 1,1)}{a_{plan}}} \\ t_p(a_{1,1}) = \sqrt{\frac{2 \cdot (l_g 2 + l_a 1,1)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \\ t_p(g_2) = \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \end{cases}$$

Puis en injectant (D.1) :

$$\Leftrightarrow \begin{cases} t_p(a_{1,2}) = \sqrt{\frac{2 \cdot (l_g 2 + l_T + l_a 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot (l_g 2 + l_T)}{a_{plan}}} \\ t_p(a_{1,1}) = \sqrt{\frac{2 \cdot (l_g 2 + l_T)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \\ t_p(g_2) = \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \end{cases}$$

Pour la phase de freinage :

$$\begin{cases} t_p(c_2 + g_{1,1} + g_{1,2}) = \sqrt{\frac{2 \cdot (l_c 2 + l_g 1,1 + l_g 1,2)}{a_{plan}}} \\ t_p(g_{1,1} + g_{1,2}) = \sqrt{\frac{2 \cdot (l_g 1,1 + l_g 1,2)}{a_{plan}}} \\ t_p(g_{1,2}) = \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \end{cases}$$

$$\Leftrightarrow \begin{cases} t_p(c_2) = \sqrt{\frac{2 \cdot (l_c 2 + l_g 1,1 + l_g 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot (l_g 1,1 + l_g 1,2)}{a_{plan}}} \\ t_p(g_{1,1}) = \sqrt{\frac{2 \cdot (l_g 1,1 + l_g 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \\ t_p(g_{1,2}) = \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \end{cases}$$

Puis en injectant (D.1) :

$$\Leftrightarrow \begin{cases} t_p(c_2) = \sqrt{\frac{2 \cdot (l_c 2 + l_T + l_g 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot (l_T + l_g 1,2)}{a_{plan}}} \\ t_p(g_{1,1}) = \sqrt{\frac{2 \cdot (l_T + l_g 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \\ t_p(g_{1,2}) = \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \end{cases}$$

De même les segments suivants sont parcourus à la vitesse constante $v_{plan\ max}$: a_2 , $b_{1,1}$, $b_{1,2}$, $c_{1,1}$ et $c_{1,2}$. Or pour un segment de longueur d parcouru à la vitesse constante $v_{plan\ max}$, on a la relation : $t_p(d) = d / v_{plan\ max}$.

On peut donc écrire le système suivant :

$$\begin{cases} t_p(a_2) = l_{a\ 2} / v_{plan\ max} \\ t_p(b_{1,1}) = l_{b\ 1,1} / v_{plan\ max} \\ t_p(b_{1,2}) = l_{b\ 1,2} / v_{plan\ max} \\ t_p(c_{1,1}) = l_{c\ 1,1} / v_{plan\ max} \\ t_p(c_{1,2}) = l_{c\ 1,2} / v_{plan\ max} \end{cases}$$

Puis en remplaçant les longueurs connues :

$$\Leftrightarrow \begin{cases} t_p(a_2) = l_{a\ 2} / v_{plan\ max} \\ t_p(b_{1,1}) = l_T / v_{plan\ max} \\ t_p(b_{1,2}) = l_{b\ 1,2} / v_{plan\ max} \\ t_p(c_{1,1}) = l_T / v_{plan\ max} \\ t_p(c_{1,2}) = l_{c\ 1,2} / v_{plan\ max} \end{cases} \quad (\text{D.6})$$

On suppose que les paramètres suivants sont fixés : t_{stop} , $v_{plan\ max}$, l_T et a_{plan} . Nous avons encore trop d'inconnues pour résoudre le système (3.2), on pose donc : $l_{g\ 2} = l_{g\ 1,2}$ valeur que l'on suppose également fixée. On rassemble toutes les expressions de $t_p(0)$ en injectant les équations (D.5). On surligne les inconnues restantes :

$$\left\{ \begin{array}{l} t_p(a_{1,1}) = \sqrt{\frac{2 \cdot (l_g 2 + l_T)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \\ t_p(a_{1,2}) = \sqrt{\frac{2 \cdot \frac{v_{plan max}^2}{2 \cdot a_{plan}}}{a_{plan}}} - \sqrt{\frac{2 \cdot (l_g 2 + l_T)}{a_{plan}}} \\ t_p(a_2) = l_a 2 / v_{plan max} \\ t_p(b_{1,1}) = l_T / v_{plan max} \\ t_p(b_{1,2}) = l_b 1,2 / v_{plan max} \\ t_p(c_{1,1}) = l_T / v_{plan max} \\ t_p(c_{1,2}) = l_c 1,2 / v_{plan max} \\ t_p(c_2) = \sqrt{\frac{2 \cdot \frac{v_{plan max}^2}{2 \cdot a_{plan}}}{a_{plan}}} - \sqrt{\frac{2 \cdot (l_T + l_g 1,2)}{a_{plan}}} \\ t_p(g_{1,1}) = \sqrt{\frac{2 \cdot (l_T + l_g 1,2)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \\ t_p(g_{1,2}) = \sqrt{\frac{2 \cdot l_g 1,2}{a_{plan}}} \\ t_{stop} = t_{stop} \\ t_p(g_2) = \sqrt{\frac{2 \cdot l_g 2}{a_{plan}}} \end{array} \right.$$

Il apparaît par ailleurs que :

$$\left\{ \begin{array}{l} t_p(a_{1,1}) = t_p(g_{1,1}) \\ t_p(a_{1,2}) = t_p(c_2) \\ t_p(b_{1,1}) = t_p(c_{1,1}) \\ t_p(g_{1,2}) = t_p(g_2) \end{array} \right. \quad (D.7)$$

Le système d'équation (3.2) devient (on surligne encore les inconnues) :

$$(3.2) \Leftrightarrow \left\{ \begin{array}{l} t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + 2 \cdot t_p(b_{1,1}) + t_p(b_{1,2}) = t \\ t_p(a_{1,1}) + t_p(a_{1,2}) + 2 \cdot t_p(b_{1,1}) + t_p(b_{1,2}) + t_p(c_{1,2}) = t \\ 2 \cdot t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(b_{1,1}) + t_p(c_{1,2}) + 2 \cdot t_p(g_{1,2}) + t_{stop} = t \\ 2 \cdot t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + t_p(b_{1,1}) + 2 \cdot t_p(g_{1,2}) + t_{stop} = t \end{array} \right.$$

La soustraction des lignes 1 et 2 d'une part et 3 et 4 d'autre part donne le même résultat¹ qu'on réinjecte dans le système à savoir :

$$t_p(a_2) = t_p(c_{1,2})$$

1. Il apparaît que le système possédait des équations liées linéairement.

$$(3.2) \Leftrightarrow \begin{cases} t_p(a_2) = t_p(c_{1,2}) \\ t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + 2 \cdot t_p(b_{1,1}) + t_p(b_{1,2}) \\ 2 \cdot t_p(a_{1,1}) + t_p(a_{1,2}) + t_p(a_2) + t_p(b_{1,1}) + 2 \cdot t_p(g_{1,2}) + t_{stop} = t \end{cases}$$

$$(3.2) \Leftrightarrow \begin{cases} t_p(a_2) = t_p(c_{1,2}) \\ t_p(b_{1,2}) = t_p(a_{1,1}) + 2 \cdot t_p(g_{1,2}) + t_{stop} - t_p(b_{1,1}) \end{cases} \quad (D.8)$$

Mais il existe encore une infinité de solutions pour satisfaire ce système d'équations à cause de l'indétermination de la seconde ligne. Il ne faut pas oublier que prendre les temps les plus petits possibles permettra d'optimiser la capacité de la ligne sous contrainte de respecter les distances de sécurité de freinage. En pratique dans ce manuscrit on lève donc cette indétermination en collant au plus près à ces distances de sécurité mais avec une petite marge. (voir en fin de cette annexe)

Il reste à déterminer les longueurs des onze segments de cantons. On sait déjà que :

$$\begin{array}{ll} l_{a,1,1} = l_{b,1,1} = l_{c,1,1} = l_g,1,1 = l_T & 4 \text{ variables fixées} \\ l_g,2 = l_g,1,2 = (\text{fixé par choix}) & 2 \text{ variables fixées} \end{array}$$

Et à partir des équations (D.7) et (D.5), on peut déterminer en plus que :

$$\begin{array}{ll} l_{a,1,2} = l_c,2 = \frac{v_{plan\ max}^2}{2 \cdot a_{plan}} - l_g,2 - l_T & 2 \text{ variables fixées} \\ l_{a,2} = l_{c,1,2} & 1 \text{ variable fixée} \end{array}$$

Enfin la seconde équation du système (D.8) peut s'écrire :

$$\begin{aligned} \frac{l_{b,1,2}}{v_{plan\ max}} &= \sqrt{\frac{2 \cdot (l_g,2 + l_T)}{a_{plan}}} - \sqrt{\frac{2 \cdot l_g,2}{a_{plan}}} + 2 \cdot \sqrt{\frac{2 \cdot l_g,1,2}{a_{plan}}} + t_{stop} - \frac{l_T}{v_{plan\ max}} \\ \Leftrightarrow l_{b,1,2} &= v_{plan\ max} \cdot \left(\sqrt{\frac{2 \cdot (l_g,2 + l_T)}{a_{plan}}} + \sqrt{\frac{2 \cdot l_g,1,2}{a_{plan}}} + t_{stop} \right) - l_T \quad 1 \text{ variable fixée} \end{aligned}$$

On a fixé dix variables sur onze, la dernière variable est fixé comme on l'a dit plus haut en levant l'indétermination du système (D.8).

Application numérique :

À partir des formules précédentes, on peut calculer les longueurs et les temps de parcours de tous les segments de cantons. Les variables ainsi que leurs valeurs prises dans le cadre du chapitre 3 sont résumées dans le tableau 3.1 au chapitre 3 aussi bien pour les variables fixées au départ que pour celles déterminées par le calcul.

