

ÉCOLE DOCTORALE SCIENCES ET MÉTIERS DE L'INGÉNIEUR

**Laboratoire d'Électrotechnique et d'Électronique de Puissance
ULR 2697 - Campus de Lille**

THÈSE

présentée par : **Luis-Maria ALONSO**
soutenue le : **26 Janvier 2022**

pour obtenir le grade de : **Docteur d'HESAM Université**

préparée à : **École Nationale Supérieure d'Arts et Métiers**
Spécialité : **Génie électrique**

**Modélisation et simulation de la consommation énergétique
des trains dans le contexte ferroviaire ouvert à la concurrence
pour un décompte d'énergie à des fins de facturation**

THÈSE dirigée par :
M. Benoît ROBYNS et M. Christophe SAUDEMONT

et co-encadrée par :
M. Laurent ROUX, M. Lionel TAUNAY et M. Aurelien WATARE

Jury

Mme. Paola PELLEGRINI
M. Nicolas RETIERE
M. Bruno SARENI
M. Benoît ROBYNS
M. Christophe SAUDEMONT
M. Laurent ROUX
M. Lionel TAUNAY

Directrice de recherche, Université Gustave Eiffel
Professeur, Université de Grenoble Alpes
Professeur, INP-ENSEEIH
Professeur, Junia
Professeur, Junia
Ingénieur, RTE
Ingénieur, SNCF Réseau

Présidente
Rapporteur
Rapporteur
Examineur
Examineur
Invité
Invité

**T
H
È
S
E**

Cette thèse a été financée par RTE et SNCF Réseau.

Remerciements

Cette thèse est le résultat de trois années de travail, dans des conditions très particulières, qui ont fait de cette thèse un travail international du fait de mes différents lieux de confinement. Je voudrais remercier toutes les personnes qui m'ont aidé d'une manière ou d'une autre à finir en temps et en heure cette thèse.

Je voudrais d'abord remercier mon équipe d'encadrants chez JUNIA ; Benoît Robyns et Christophe Saudemont pour m'avoir guidé dans le complexe monde des publications scientifiques. De même, sans l'aide et les connaissances du domaine ferroviaire apportées par Lionel Taunay et Étienne Sourdille de SNCF Réseau, cette thèse ne serait pas arrivée à bon port. Finalement, je voudrais aussi envoyer un grand merci à Laurent Roux, ferrovipathe comme moi, qui depuis RTE a lancé le projet en partenariat avec SNCF Réseau pour donner naissance à ma thèse. Ton implication dans cette thèse, même quand cela ne faisait plus parti de tes obligations, a été clé pour pouvoir aboutir.

Je souhaite ensuite remercier tous les membres du jury pour avoir accepté de pendre et d'assister en personne à la soutenance de ma thèse malgré le contexte sanitaire. Cela a aidé à réduire le stress propre d'une soutenance et m'a épargné quelques prières aux dieux de l'informatique et des téléconférences.

De façon plus personnelle, je tiens à remercier mes très chers amis sans qui mon passage par JUNIA n'aurait pas été le même. Je dois dire merci à Abderrahman, pour être l'explorateur qui m'a averti de toutes les pièges, à Anthony (le stagiaire qui a réussi à monter en grade), pour ses commentaires random qui étaient le point de départ pour des discussions interminables, à Yacine (même si j'ai mal prononcé son prénom de façon continu et inconsciente), pour les moments de convivialité partagés à JUNIA et ailleurs, à Matthieu, pour les discussions sur la droite et le libéralisme et à Adrien, même si nous n'avons passé que peu de temps ensemble, j'ai apprécié nos discussions sur la gauche et sur le sens de la vie. Et bien-sûr, je ne peux pas oublier le *haut conseil des grands docteurs* qui étant composé exclusivement de membres JUNIA, avait son siège dans les bars de Lille. Makram, Jad et Benoît, vous êtes tous les trois très différents mais l'approche personnelle que vous avez est la même et elle m'a beaucoup aidé au cours de mes séjours à Lille. Je veux aussi saluer tous les doctorants,

REMERCIEMENTS

post-doctorants, professeurs et stagiaires que j'ai pu croiser lors de mon passage chez JUNIA. Des pauses café au repas qu'on a pu partager ensemble resteront toujours dans ma mémoire. Je voudrais aussi remercier tous mes amis français de Toulouse ou Rennes qui ont égayé mon séjour en France et m'ont donné envie de poursuivre mon séjour de double diplôme en thèse.

Y no podría acabar esta sección sin agradecer a mi familia y amigos españoles su apoyo. A ellos, que siempre han estado al lado mío en los buenos y en los malos momentos, también les dedico esta tesis.

REMERCIEMENTS

Résumé

A travers différentes directives, l'Union Européenne définit les conditions et modalités de l'ouverture à la concurrence du transport ferroviaire et du marché de l'électricité. Ainsi, les Entreprises Ferroviaires (EF) qui affrètent des trains circulants sur le réseau ferré Français, peuvent s'approvisionner en énergie électrique auprès du fournisseur de leur choix. La juste affectation aux EF, de l'énergie consommée par leur flotte d'engins nécessite des compteurs d'énergie embarqués, communiquant vers un système de collecte et de facturation.

Face aux enjeux financiers et aux enjeux d'équilibrage du système électrique, RTE et SNCF Réseau doivent garantir une bonne fiabilité du comptage de l'énergie aux EF(s), malgré les dysfonctionnements des compteurs qui ne permettent pas l'exhaustivité des mesures au service d'une consommation effective.

Pour répondre à la problématique de la fiabilité des compteurs, des bornes de validation des consommations relevées par les compteurs basés sur des modèles physiques ont été développées. Ces bornes permettent d'augmenter la fiabilité des données de comptage en écartant les mesures aberrantes du point de vue physique.

Afin d'améliorer la disponibilité des données de comptage, des estimateurs sont développés dans cette thèse avec comme objectif d'estimer la consommation d'un train pour une circulation donnée, en utilisant une base de données composée par des mesures de compteurs, réalisées sur une ou plusieurs années, validées par les bornes définies précédemment. Les différents estimateurs utilisent en plus diverses sources d'information (temporelles, météo, consommations mesurées dans des sous-stations d'alimentation,...) afin d'améliorer l'estimation de la consommation. Les sources d'information les plus pertinentes sont identifiées, permettant ainsi d'éliminer les sources de bruit.

L'impact de l'évolution du réseau ferroviaire vers un smart grid ferroviaire sur la consommation ferroviaire et le comptage de l'énergie est abordé en fin de thèse.

Mots-clés : Comptage énergétique, ferroviaire, validation de la consommation, estimation de la consommation.

RESUME

Abstract

Via various directives, the European Union is progressively opening up the rail transport and electricity markets to competition in Europe. As a result, railway undertakings (RUs) that operate trains on the French rail network can obtain their electrical energy from the supplier of their choice. The correct allocation to RUs of the energy consumed by their fleet of vehicles requires on-board energy meters that communicate with a data collection and billing system.

In view of the financial stakes, RTE and SNCF Réseau must guarantee reliable energy metering to RUs, despite energy meter malfunctions that do not allow for exhaustive measurements of the actual consumption.

In order to address the issue of the energy meters' reliability, validation bounds for the consumption read by the energy meters based on physical models have been developed. These bounds increase the reliability of the metering data by eliminating outliers from the physical point of view.

In order to improve the availability of metering data, estimators are developed in this thesis with the objective of estimating the consumption of a train for a given journey, using a database composed of energy meter measurements, performed over one or more years, validated by the previously defined bounds. The various estimators also use various sources of information (time, weather, consumption measured in supply substations, etc.) in order to improve the consumption estimation. The most relevant sources of information are identified, allowing to eliminate sources of noise.

The impact of the evolution of the railway network towards a railway smart grid on railway consumption and energy metering is addressed at the end of the thesis.

Keywords : Energy metering, railways, consumption validation, consumption estimation.

ABSTRACT

Table des matières

Remerciements	5
Résumé	9
Abstract	11
Liste des tableaux	18
Liste des figures	21
Liste des acronymes et notations	22
Introduction générale	23
1 Validation des mesures des compteurs : génération des bornes de validation	25
1.1 Introduction	26
1.2 Modèles physiques des trains dans la littérature	27
1.3 Contrôle optimal, calcul de la borne de validation inférieure	30
1.3.1 Modélisation mathématique du problème	30
1.3.2 Résolution numérique du problème de « commande optimale » avec MATLAB	31
1.4 Maximum «ferroviaire», calcul de la borne de validation supérieure	35
1.5 Calcul des bornes maximales et minimales avec un temps de trajet variable	40
1.6 Application à des cas d'étude	43
1.6.1 Cas d'étude TER n°1 Bordeaux-Arcachon	43
1.6.2 Cas d'étude TER n°2 Bordeaux-Dax	46

TABLE DES MATIÈRES

1.6.3	Cas d'étude TER n°3 Dijon-Besançon	50
1.7	Conclusion	54
2	Estimateurs de la consommation pour un trajet donné	55
2.1	Introduction	56
2.2	Estimateurs « data-based »	57
2.2.1	Moyenne sur l'ensemble de la base de données	58
2.2.2	Moyenne mensuelle	58
2.2.3	Régression linéaire multiple (RLM)	58
2.2.4	Support Vector Machine (SVM)	58
2.2.5	Réseaux de neurones	59
2.3	Méthode pour trier les sources d'information disponibles	59
2.4	Cas d'étude TER	60
2.4.1	Sources d'information disponibles	60
2.4.2	Cas d'étude TER n°1 Bordeaux-Arcachon	61
2.4.3	Cas d'étude TER n°2 Bordeaux-Dax	65
2.4.4	Cas d'étude TER n°3 Dijon-Besançon	66
2.5	Cas d'étude FRET	68
2.5.1	Sources d'information disponibles	68
2.5.2	Cas d'étude FRET n°1 Avignon-Valenton	69
2.5.3	Cas d'étude FRET n°2 Le Boulou - Zoufftgen	71
2.5.4	Cas d'étude FRET n°3 Dunkerque - Ebange	74
2.6	Conclusion	76
3	Impact des smart-grids ferroviaires dans le comptage d'énergie	79
3.1	Introduction	80
3.2	Impact des smart-grids dans la consommation des trains	80
3.2.1	Analogie entre les électrifications en DC et AC vis-à-vis de l'énergie récupérée .	81
3.2.2	Effacement de la consommation ferroviaire sur demande	88
3.3	Impact des smart-grids dans les estimateurs de la consommation	89

TABLE DES MATIÈRES

3.4 Conclusion	90
Conclusion et perspectives	91
Liste des publications	95
Bibliographie	97
Annexe 1	103
Annexe 2	107

TABLE DES MATIÈRES

Liste des tableaux

1.1	Impact de l'algorithme sur la précision et le temps de calcul	34
1.2	Impact du pas de calcul sur la précision et le temps de calcul avec l'algorithme SQP .	34
1.3	Impact du pas de discrétisation h sur la précision et le temps de calcul avec l'algorithme SQP	34
1.4	Temps de parcours entre les arrêts de la ligne Bordeaux-Arcachon	44
1.5	Temps de parcours entre les arrêts de la ligne Bordeaux-Dax	48
1.6	Temps de parcours entre les arrêts de la ligne Dijon-Besançon	51
2.1	Incertitude du modèle physique liée à la masse totale du train	56
2.2	Incertitude du modèle physique liée au vent	57
2.3	Performance des estimateurs de la consommation pour le cas d'étude Bordeaux-Arcachon	62
2.4	Temps de calcul pour la generation chaque estimateur pour le cas d'étude Bordeaux-Arcachon	62
2.5	Apport de chaque source d'information pour l'estimateur SVM pour le cas d'étude Bordeaux-Arcachon	63
2.6	Comparaison des performances des estimateurs avec information des sous-stations toutes les 10 min et 1 min pour le cas d'étude Bordeaux-Arcachon 2018	64
2.7	Comparaison de l'apport de chaque source d'information pour l'estimateur SVM avec information des sous-stations toutes les 10 min et 1 min pour le cas d'étude Bordeaux-Arcachon 2018	64
2.8	Performance des estimateurs de la consommation pour le cas d'étude Bordeaux-Dax .	65
2.9	Apport de chaque source d'information pour l'estimateur SVM pour le cas d'étude Bordeaux-Dax	66

LISTE DES TABLEAUX

2.10	Performance des estimateurs de la consommation pour le cas d'étude Dijon-Besançon	67
2.11	Apport de chaque source d'information pour l'estimateur SVM pour le cas d'étude Dijon-Besançon	68
2.12	Performance des estimateurs de la consommation pour le cas d'étude Avignon-Valenton	70
2.13	Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Avignon-Valenton	71
2.14	Performance des estimateurs de la consommation pour le cas d'étude Dunkerque-Ebange	72
2.15	Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Le Boulou-Zoufftgen	73
2.16	Performance des estimateurs de la consommation pour le cas d'étude Dunkerque-Ebange	75
2.17	Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Dunkerque-Ebange	75
2.18	Apport moyen de chaque source d'information pour l'estimateur SVM pour les cas d'étude TER	76
2.19	Apport moyen de chaque source d'information pour le meilleur estimateur pour les cas d'étude FRET	77
3.1	Comparaison des performances des estimateurs dans les cas d'étude FRET avec et sans les consommations dans les sous-stations comme source d'information	89
A2.3	Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Avignon-Valenton	110
A2.5	Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Le Boulou-Zoufftgen	119
A2.6	Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Dunkerque-Ebange	125

Liste des figures

1.1	Représentation théorique des bornes maximales et minimales avec différentes mesures relevées	27
1.2	Composition d'un essieu ferroviaire, les boudins sont une partie du système de guidage ferroviaire [3]	28
1.3	« Efficiency map » d'un moteur asynchrone [11]	29
1.4	Consommation minimale (en vert), consommation maximale (en rouge) et à vitesse maximale (en bleu) pour un trajet donné	36
1.5	Profil de vitesse et commande pour une consommation maximale dans le sens mathématique	37
1.6	Puissance développée par le train dans toute sa plage de vitesse	37
1.7	Profil de vitesse et de force pour un trajet avec une consommation maximale « ferroviaire »	38
1.7	Calcul du point de freinage avec l'inversion de modèle	39
1.8	Consommation minimale (en vert), à vitesse maximale (en bleu), consommation maximale « ferroviaire » (en orange) et consommation maximale « mathématique » (en rouge) pour un trajet donné	39
1.9	Comparaison entre le minimum calculé par la force brute et par optimisation pour un trajet de 8 arrêts	41
1.10	Bornes maximales et minimales pour un trajet Bordeaux-Arcachon avec différents temps de parcours	42
1.11	Situation géographique de la ligne étudié	44
1.12	Bornes de validation inférieures pour les deux sens de circulation de la ligne Bordeaux-Arcachon	45
1.13	Bornes de validation supérieures pour le deux sens de circualtion de la ligne Bordeaux-Arcachon	45

LISTE DES FIGURES

1.14	Bornes de validation supérieures pour les deux sens de circulation de la ligne Bordeaux-Arcachon avec les mesures des compteurs pour les années 2017, 2018 et 2019	46
1.15	Situation géographiquet de la ligne étudié	47
1.16	Bornes de validation supérieures pour les deux sens de circulation de la ligne Bordeaux-Dax avec les mesures des compteurs pour les années 2017, 2018 et 2019	48
1.17	Bornes de validation supérieures pour la ligne Bordeaux-Dax avec les mesures des compteurs pour le sens de circulation Bordeaux-Dax pour les années 2017, 2018 et 2019 . .	49
1.18	Bornes de validation supérieures pour la ligne Bordeaux-Dax avec les mesures des compteurs pour le sens de circulation Dax-Bordeaux les années 2017, 2018 et 2019	50
1.19	Situation géographiquet de la ligne étudié	51
1.20	Bornes de validation inférieures pour les deux sens de circulation de la ligne Dijon-Besançon	52
1.21	Bornes de validation supérieures pour les deux sens de circulation de la ligne Dijon-Besançon	52
1.22	Bornes de validation supérieures pour les deux sens de circulation de la ligne Dijon-Besançon avec les mesures des compteurs pour les années 2017, 2018 et 2019	53
2.1	Topologie de reseau de neurones utilisée dans cette thèse [36]	59
2.2	Trajet Avignon-Valenton avec les différents parcours	70
2.3	Trajet Le Boulou - Zoufftgen avec les différents parcours	72
2.4	Trajet Dunkerque - Ebange avec les différents parcours	74
3.1	Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation DC	82
3.2	Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation AC	83
3.3	Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets Dijon-Besançon	84
3.4	Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets Besançon-Dijon	84
3.5	Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation DC	85

3.6	Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation AC	86
3.7	Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets Le Boulou – Zoufftgen	87
3.8	Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets Zoufftgen-Le Boulou	87
3.9	Différents profils de vitesse pour un trajet donné sans effacement en rouge, avec effacement et sans retard en bleu et avec effacement et retard en vert [46]	88

Liste des acronymes et notations

α	Coefficient de masses tournantes
μ	Facteur qui varie entre 1 et 0 qui reflète la capacité de l'infrastructure à accepter l'énergie de freinage récupérée par un train
θ	L'angle de la pente de la ligne
F_g	Force due à la déclivité
g	La gravité
h	Pas de discrétisation
M	Masse du train
R_c	Rayon de courbure [m]
r_c	Resistance en courbe par tonne
<i>s.l.c.</i>	Sous les contraintes
CVS	Convertisseur statique
EF	Entreprise Ferroviaire
FRET	Transport de marchandise
PMP	Principe du Maximun de Pontryagin
RLM	Régression linéaire Multiple
SVM	Support Vector Machine
TBC	Tonnage Brut Convoi, Tonnage Brut Remorqué plus tonnage des motrices
TBR	Tonnage Brut Remorqué
TER	Transport Express Régional
UM	Unité multiple, plusieurs machines pour un même train

Introduction générale

Contexte

À travers une série de directives, l'Union Européenne définit les conditions et modalités de l'ouverture à la concurrence du transport ferroviaire et du marché de l'électricité. Ainsi, les Entreprises Ferroviaires (EF) qui opèrent sur le réseau ferré français peuvent s'approvisionner en énergie électrique auprès du fournisseur de leur choix pour l'alimentation de leurs locomotives et automotrices. La puissance de chaque engin de traction est de plusieurs mégawatts. De telles puissances autorisent les EF à s'approvisionner sur les marchés de gros à des prix attractifs. Ces importants volumes de consommation sont par ailleurs rattachés au périmètre des responsables d'équilibre désignés par les EF. Le dispositif de Responsable d'Équilibre est un dispositif européen défini par le RÈGLEMENT (UE) 2017/2195 DE LA COMMISSION du 23 novembre 2017 concernant une ligne directrice sur l'équilibrage du système électrique. Face à de tels enjeux, financiers d'une part, et relatifs à l'équilibrage des systèmes électriques d'autre part, RTE (désigné par la loi française comme le responsable du comptage d'énergie de chaque EF) et SNCF Réseau (en charge du maintien et de la gestion du réseau ferroviaire français) doivent garantir une bonne fiabilité de la mesure de l'énergie consommées par chaque EF, pour la facturation de l'énergie réelle qu'ils consomment. Lorsque SNCF était un monopole, il n'y avait qu'un seul consommateur d'énergie et les compteurs étaient installés à la frontière entre EDF et SNCF c'est à dire dans les sous stations, puis avec l'ouverture à la concurrence la frontière électrique s'est déplacée au niveau des locomotives. Il a fallu installer des compteurs dans les engins. Or vu de RTE (successeur d'EDF), les compteurs installés dans les sous-stations mesurent l'ensemble de l'activité ferroviaire. Ainsi l'objectif du décompte des énergies consiste à attribuer correctement cette énergie mesurée à l'ensemble des acteurs SNCF Réseau et les EF. Maîtriser l'énergie consommée par chaque EF pour la facturation a un intérêt au niveau européen (des projets sont menés en Europe sur la même thématique [1]), donc pour essayer de remédier le problème du comptage individuel de chaque EF, l'Union Européenne prévoit l'installation obligatoire de compteurs dans le matériel roulant construit ou rénové depuis 2014. Malheureusement, ces compteurs n'ont pas montré une fidélité de mesure complète en donnant parfois des mesures aberrantes. En même temps, ces compteurs renvoient l'information à travers les réseaux de téléphonie qui parfois n'ont pas la meilleure couverture pour

toute les lignes ferroviaires. Cela fait que les relevés sont ainsi parfois mal transmis ou transmis avec beaucoup de retard.

SNCF Réseau dispose de logiciels historiques comme ESMERALDA (Étude par Simulation sur Modèles Électriques Représentatifs de l’Alimentation et des Locomotives pour le Dimensionnement des Appareils) qui permet de calculer l’énergie consommée par un train dans une marche type. Même si ces calculs ne sont valables pour le dimensionnement des appareils de l’infrastructure électrique de traction (section de caténaire, transformateurs, redresseurs...), le logiciel ESMERALDA ne permet pas d’estimer la consommation d’un train donné pour un trajet donnée. Si d’autres logiciels ont été développés récemment, comme le simulateur OPEUS du projet européen Shift2Rail, cet outil n’est pas adapté pour répondre à la problématique spécifique de cette thèse. Le simulateur OPEUS a pour but d’analyser les performances des différentes parties de la chaîne de traction d’un train à partir de cas d’études concrets.

Objectifs

Cette thèse a pour objectif de développer un système capable d’augmenter la fiabilité des mesures relevées par les compteurs embarqués dans le train et d’améliorer la disponibilité des données de comptage en estimant la consommation quand la mesure n’est pas disponible ou fiable.

Dans le premier chapitre, une méthode de génération de bornes de validation basée sur des modèles physiques est présentée afin d’augmenter la fiabilité des mesures relevées par les compteurs embarqués. Les modèles physiques des trains décrits dans la littérature seront évoqués, le “contrôle optimal” utilisé pour déterminer la consommation minimale atteignable sera présenté ainsi que la notion du maximum “ferroviaire” servant à calculer la consommation maximale atteignable. En fin de chapitre, cette méthode sera appliquée à des cas d’étude TER.

Dans le deuxième chapitre, afin d’améliorer la disponibilité des données de comptage divers estimateurs de la consommation des trains seront évalués. Ces derniers utiliseront plusieurs sources d’information qui, selon le cas d’étude, peuvent devenir sources de bruit. Une classification par importance des sources et la performance de chaque estimateur est établie pour des cas d’étude TER et FRET.

Dans le troisième chapitre, une réflexion sur l’impact des smart-grids ferroviaires dans la structure de comptage est menée ainsi que son possible impact dans les consommations totales des trains.

Le dernier chapitre offre un résumé des résultats obtenus et envisage des perspectives.

Chapitre 1

Validation des mesures des compteurs : génération des bornes de validation

1.1 Introduction

RTE a été désigné par le code de l'énergie français comme responsable des activités de comptage nécessaires à l'exercice de ses missions. A ce titre, RTE est responsable du comptage et du décompte des énergies de chaque EF consommatrice sur le réseau ferré français. Notons que RTE possède déjà une longue expérience du comptage d'énergie pour ses clients classiques raccordés directement ou indirectement à son réseau public de transport où deux compteurs sont installés en parallèle pour assurer l'intégrité des mesures. Dans le cas du comptage d'énergie dans le domaine ferroviaire, cela n'est pas le standard, la loi européenne ne prévoyant que l'installation d'un seul compteur. Afin d'obtenir le même niveau de fidélité dans le comptage d'énergie ferroviaire que dans le comptage d'énergie industriel, RTE souhaite développer une méthode permettant de valider les mesures relevées par les compteurs.

Dans ce chapitre, une méthode pour générer des bornes de validation basées sur des modèles physiques sera présentée. Ces bornes représentent les consommations atteignables du point de vue physique, ce qui permet d'identifier les consommations aberrantes, et donc, les compteurs en panne. Comme illustré dans la Figure 1.1, la méthode développée permet de générer une valeur maximale et minimale pour différents temps de trajet qui permettent de valider ou non la mesure relevée par le compteur embarqué. Les modèles physiques abordés par la littérature seront ainsi passés en revue avant de décrire les modélisations faites pour obtenir les bornes inférieure et supérieure. Les incertitudes liées au modèle physique seront exposées ainsi que la solution adoptée pour prévoir des bornes de validation pour des temps de trajets différents. Finalement, cette méthode sera appliquée à un cas d'étude de type TER.

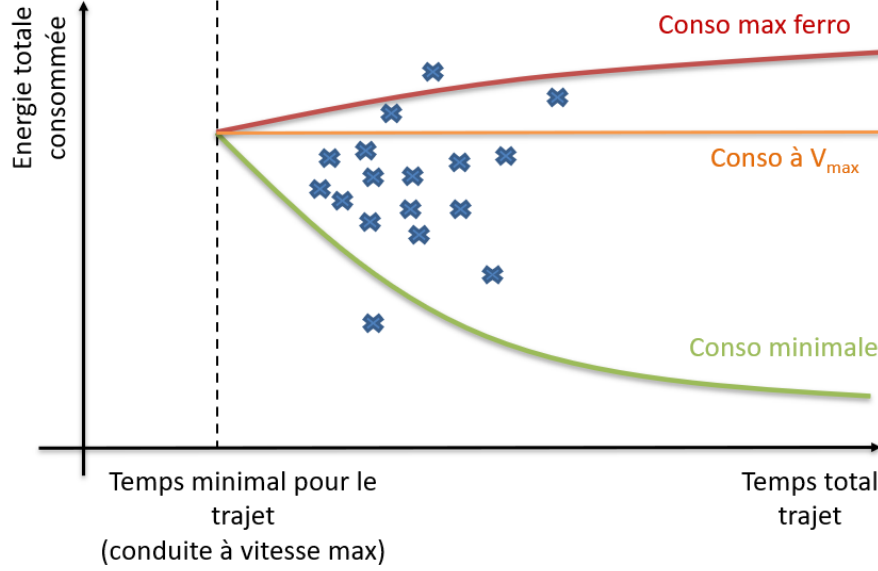


FIGURE 1.1 – Représentation théorique des bornes maximales et minimales avec différentes mesures relevées

1.2 Modèles physiques des trains dans la littérature

La modélisation de la consommation énergétique des trains est un sujet largement traité dans la littérature ferroviaire [2]. Dans tous les modèles existants, trois forces majeures sont mentionnées : la force liée à la déclivité de la ligne, la force en courbe produite pour les frottements supplémentaires des boudins dans les rails, et la force de résistance à l'avancement qui tient compte de tous les frottements aérodynamiques et mécaniques présents dans le train.

Force due à la déclivité (F_g) : quand le train est dans une pente, on a une force qui s'oppose à l'avancement du train. Dans la littérature, cette force se modélise toujours de la façon suivante :

$$F_g = M * g * \sin(\theta) \quad (1.1)$$

Avec M la masse du train, g la gravité et θ l'angle de la pente de la ligne. Comme les pentes présentes dans les lignes ferroviaires sont très faibles (de l'ordre de quelques ‰) le terme $\sin(\theta)$ peut être approximé par θ directement pour simplifier les calculs :

$$F_g = M * g * \theta \quad (1.2)$$

La force en courbe (F_c) : quand le train entre dans une courbe, les efforts roues/rails augmentent en générant une force résistante plus grande liée aux frottements supplémentaires des boudins dans les courbes.

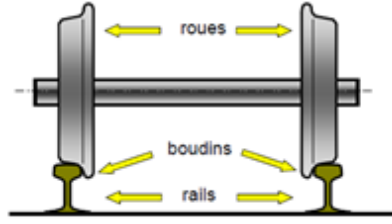


FIGURE 1.2 – Composition d'un essieu ferroviaire, les boudins sont une partie du système de guidage ferroviaire [3]

Plusieurs formules sont proposées dans la littérature pour modéliser cette force qui ont été résumées dans l'annexe 1. Dans cette thèse la formule 1.3, proposée par la SNCF dans [4], a été retenue étant une des plus simples pour modéliser cette force.

$$r_c[kN/t] = \frac{g * \sin\left(\frac{800}{R_c}\right)}{1000} \quad (1.3)$$

La force de résistance à l'avancement (F_{res}) : Cette force comprend toutes les forces qui empêchent le mouvement du train, comme les forces du frottement entre les roues et les rails, la force de trainage aérodynamique du train, la force du frottement du pantographe, etc. Dans la littérature, cette force est représentée à travers des fonctions quadratiques en fonction de la vitesse du type suivant :

$$F_{res} = A + B * V + C * V^2 \quad (1.4)$$

Avec A, B et C des coefficients liés à des dimensions caractéristiques de chaque train et V la vitesse linéaire du train. La détermination des coefficients A, B et C a été l'objet de nombreux travaux, mais on distingue deux tendances principales pour les calculer : soit on les reconstruit depuis des essais [5][6], soit on les estime à partir de modèles semi-empiriques. Dans la littérature, plusieurs propositions de modèles sont disponibles [7] ; elles sont présentées dans l'annexe 1.

Les valeurs des coefficients A, B, C sont connues par l'entreprise en charge des infrastructures [8]. La transmission de ces valeurs entre autres est essentielle pour homologuer le matériel. Dans cette thèse, les coefficients A, B et C du matériel roulant étudié sont fournies par SNCF Réseau.

Des facteurs de corrections sont aussi proposés pour prendre en compte les circonstances climatiques (vent, température, humidité, etc.) d'exploitation [9]. Ces circonstances sont prises en compte dans le terme $C * V^2$:

$$\rho * C * (V_{train} + V_{vent})^2 \quad (1.5)$$

Dans ρ (densité de l'air), on prendra en compte les conditions d'humidité et de température ambiantes. En rajoutant V_{vent} (vitesse du vent) à la vitesse du train, on pourra calculer la vitesse apparente

de l'air qui freine le train. On peut aussi trouver des facteurs de correction pour tenir compte de la résistance additionnelle lors de la traversée d'un tunnel.

Dans la chaîne de traction, on peut aussi trouver des consommations dites « auxiliaires » qui comportent les consommations des équipements de la machine (ventilateurs, systèmes de sécurité, chargeurs des batteries, etc.) et les consommations pour le confort des voyageurs (climatisation, chauffage, illumination, etc.). Pour modéliser ces consommations, deux approches sont envisageables : soit on suppose des valeurs constantes considérées « habituelles » soit on utilise des formules pour estimer la variation de puissance de ces auxiliaires avec la température extérieure, le nombre d'arrêts, etc.[10]

Finalement, pour estimer globalement l'énergie mesurée au niveau du pantographe, on doit prendre en compte les rendements des différents systèmes qui constituent la chaîne de traction (réducteur mécanique, moteur électrique, CVS, transformateur, etc.). Pour estimer toutes ces pertes, deux options sont possibles : soit on « assigne » une valeur moyenne de rendement à chaque élément de la chaîne de traction, soit on utilise une « efficiency map » (un tableau avec des rendements prédéfinis selon les vitesses et les efforts de traction comme illustré à la Figure 1.3) de chaque élément ou de l'ensemble de la chaîne de traction.

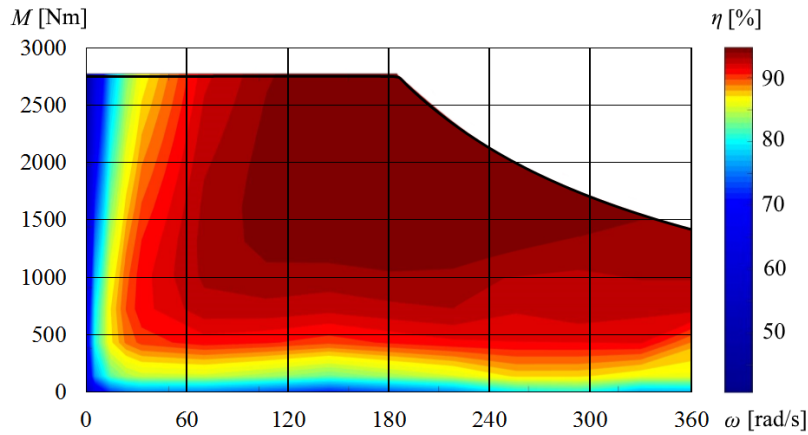


FIGURE 1.3 – « Efficiency map » d'un moteur asynchrone [11]

La supposition d'un rendement constant pour toute la plage d'effort et vitesse est discutée dans [11]. Dans le cas où la « marche sur l'erre » est plutôt favorisée face à une vitesse de croisière constante, le calcul avec un coefficient constant pour le rendement de la chaîne de traction est correct. En revanche, si la vitesse de croisière est maintenue constante, le calcul avec un coefficient constant est moins précis. Pour maintenir la vitesse de croisière, des efforts faibles sont nécessaires et font travailler la chaîne de traction dans des zones moins performantes, ce faible rendement est normalement loin du rendement constant retenu d'habitude.

1.3 Contrôle optimal, calcul de la borne de validation inférieure

1.3.1 Modélisation mathématique du problème

Pour la détermination de la borne inférieure de validation des données récupérées, la consommation minimale atteignable pour un trajet donné est recherchée, est un sujet largement traité dans la littérature sous le nom d'« eco-conduite » ou de « conduite optimale ». Plusieurs approches mathématiques y sont proposées pour calculer une conduite optimale [12][13][14] et maintenir une consommation minimale tout en respectant un horaire. La « programmation dynamique » [15][16][17] est mentionnée mais elle nécessite de grands temps de calcul en comparaison avec les autres méthodes. Une autre option est d'utiliser des méthodes dites « heuristiques » [18][19][20], mais celles-ci ont encore plus besoin de puissance de calcul. La méthode la plus utilisée est la « commande optimale » qui se base sur le principe du maximum de Pontryagin (PMP). Dans la commande optimale, la fonction objectif à minimiser est l'intégrale d'une fonction qui dépend des variables d'état. En rajoutant des contraintes dans l'optimisation, toutes les caractéristiques liées à la ligne parcourue (déclivité, rayon de courbure, etc.) par le train, ainsi que des limitations propres au train qu'on modélise (effort de traction, vitesse maximale, etc.) peuvent être prises en compte[21].

Dans le cas de la thèse, la fonction objectif est l'intégrale de soit la puissance quand la variable sur la quelle on intègre est le temps, soit une image de la puissance (en l'occurrence l'effort de traction), qui, intégrée dans le temps donne une énergie consommée comme résultat. Dans l'équation (1.6), la modélisation basée dans le temps [22] est utilisée et les deux premières contraintes sont liées directement à des équations physiques qui décrivent le mouvement du train. La troisième contrainte est liée à la capacité de traction ainsi qu'aux limitations en puissance du train. La quatrième contrainte sert à représenter les limitations de vitesse de la ligne et du train. Pour finir, les quatre dernières contraintes servent à fixer la distance à parcourir dans un temps fixé pour respecter un horaire.

$$\begin{aligned}
 \min \quad & \int_0^{t_f} [F_{trac}(t)^+ - \mu \cdot F_{trac}(t)^-] \cdot v(t) dt \\
 \text{s.l.c.} \quad & \dot{x} = v(t) \\
 & \dot{v} = \frac{F_{trac}(t) - F_{res}(v(t)) - F_g(x(t)) - F_{courbe}(x(t))}{M \cdot \alpha} \\
 & F_{trac}(t) \in [F_{trac_{min}}(v(t)), F_{trac_{max}}(v(t))] \\
 & v(t) \in [0, V_{max}(x(t))] \\
 & X(0) = 0 \quad X(t_f) = X_f \\
 & v(0) = 0 \quad v(t_f) = 0
 \end{aligned} \tag{1.6}$$

Avec F_{trac}^+ la force de traction positive, F_{trac}^- la force de traction négative (lors des freinages), μ un facteur qui reflète la capacité de l'infrastructure à accepter l'énergie récupérée, v la vitesse du train,

1.3. CONTRÔLE OPTIMAL, CALCUL DE LA BORNE DE VALIDATION INFÉRIEURE

x la position du train, F_{res} la somme de toutes les forces résistives qui s'opposent à l'avancement du train modélisée par les coefficients A, B et C, F_g la force due à la déclivité qui fait accélérer/décélérer les trains dans une pente, F_{courbe} la force supplémentaire à vaincre liée au rayon de courbure de la voie, M la masse du train, α le coefficient de masses tournantes, $F_{trac_{min}}$ la force de traction minimale qui peut être appliquée, $F_{trac_{max}}$ la force de traction maximale qui peut être appliquée et V_{max} la vitesse maximale de la ligne ou du matériel.

L'équation (1.7) représente une autre modélisation basée sur la distance et non le temps proposée par [23][24][25]. Il s'agit de la même modélisation que (1.6), mais avec un changement de variable de t pour x . Cette modélisation a l'avantage de simplifier la détermination de la contrainte de vitesse maximale et de la force additionnelle liée aux courbes en comparaison avec la modélisation basé sur le temps (la vitesse maximale seulement de x et non de $x(t)$). Néanmoins, cette modélisation a un grand désavantage, elle ne permet pas de traiter une vitesse nulle à cause de sa présence dans le dénominateur de deux des contraintes.

$$\begin{aligned}
 \min \quad & \int_{X_0}^{X_f} [F_{trac}(x)^+ - \mu \cdot F_{trac}(x)^-] dx \\
 \text{s.l.c.} \quad & \dot{t} = \frac{1}{v(x)} \\
 & \dot{v} = \frac{F_{trac}(x) - F_{res}(v(x)) - F_g(x) - F_{courbe}(x)}{M \cdot \alpha \cdot v(x)} \\
 & F_{trac}(x) \in [F_{trac_{min}}(v(x)), F_{trac_{max}}(v(x))] \\
 & v(x) \in [0, V_{max}(x)] \\
 & t(X_0) = 0 \quad t(X_f) = T_f \\
 & v(X_0) = 0 \quad v(X_f) = 0
 \end{aligned} \tag{1.7}$$

1.3.2 Résolution numérique du problème de « commande optimale » avec MATLAB

Pour solutionner les problèmes de « commande optimale », plusieurs méthodes sont proposées dans la littérature [12]. Le problème peut être résolu à travers les dites « co-variables » pour résoudre les conditions d'optimalité du PMP en résolvant des équations différentielles. Dans ce cas, la résolution est du type « indirecte » [26]. Le problème peut être aussi résolu d'une façon « directe », en étant transformé en un problème d'optimisation non linéaire. Cette dernière option a été retenue pour résoudre le problème avec le logiciel MATLAB (avec la fonction *fmincon*). Ce type de résolution est une des plus faciles à mettre en place et en même temps permet d'avoir plus de liberté sur la détermination des contraintes. Le problème (1.6) peut être écrit de la façon suivante pour le résoudre avec *fmincon* sous MATLAB :

Fonction objectif :

$$\sum_{i=1}^N (F_{trac_i}^+ - \mu \cdot F_{trac_i}^-) \cdot v_i$$

Contraintes de bornes :

$$X_0 \leq x_i \leq X_f$$

$$0 \leq v_i \leq v_{max}$$

$$F_{trac_{min}} \leq F_{trac_i} \leq F_{trac_{max}}$$

Contraintes :

$$x_{i+1} = x_i + h \cdot v_i$$

$$v_{i+1} = v_i + h \cdot \frac{F_{trac_i} - F_{res_i}(v_i)}{M \cdot \alpha}$$

$$|F_{trac_i} \cdot v_i| \leq P_{max}$$

$$x_N = X_f$$

$$v_N = V_f$$

Dans ce cas, plusieurs hypothèses simplificatrices ont dû être appliquées pour faire converger l'algorithme :

- La force additionnelle des courbes (F_c) et la force due à la déclivité (F_g) est sortie des contraintes d'optimisation et l'énergie consommée en liaison avec cette forces rajoutée dans le post-traitement
- Une seule vitesse maximale peut être imposée pour le trajet en vue de simplifier les calculs
- La limitation de l'effort de traction est effectuée à travers une limitation de puissance et non à travers la table fournie avec les spécifications du matériel roulant

Le problème (1.7) peut être écrit de la façon suivante pour le résoudre avec *fmincon* sous MATLAB :

Fonction objectif :

$$\sum_{i=1}^N (F_{trac_i}^+ - \mu \cdot F_{trac_i}^-)$$

Contraintes de bornes :

$$t_0 \leq t_i \leq t_f$$

$$0 \leq v_i \leq v_{max}$$

$$F_{trac_{min}} \leq F_{trac_i} \leq F_{trac_{max}}$$

Contraintes :

$$t_{i+1} = t_i + \frac{h}{v_i}$$

$$v_{i+1} = v_i + h \cdot \frac{F_{trac_i} - F_{res_i}(v_i) - F_{g_i} - F_{courbe_i}}{M \cdot \alpha \cdot v_i}$$

$$|F_{trac_i} \cdot v_i| \leq P_{max}$$

$$t_N = t_f$$

$$v_N = V_f$$

Dans ce cas, la formulation en termes de distance permet l'addition des forces résistives liées aux courbes et les pentes sans affecter la convergence de l'algorithme. De plus, la modélisation basée sur le temps présente une convergence (une solution au problème est trouvée) largement supérieure à la modélisation basée sur la distance, raison pour laquelle la modélisation basée sur le temps est retenue dans la suite de la thèse.

1.3.2.1 Options et temps de calcul de la fonction *fmincon* de MATLAB

La fonction *fmincon* de MATLAB possède plusieurs paramètres qui peuvent être utilisés pour fixer la précision des calculs ainsi que le temps de calcul. Dans *fmincon*, un de ces paramètres est l'algorithme pour résoudre le problème d'optimisation. Un autre point important est de fixer le pas de discrétisation h du problème : un pas très grand diminue la précision du résultat tandis qu'un pas très petit fait augmenter le nombre de variables à optimiser par la fonction très rapidement en compliquant la résolution du problème. Dans la fonction *fmincon*, les algorithmes suivants sont à disposition : « interior-point », « SQP » et « active-set ». Pour la résolution du problème basé dans le temps, avec un pas de discrétisation de 2 secondes et une tolérance sur le pas de calcul de 10^{-4} (choisis comme un point de départ fonctionnel) pour le cas d'étude du TER Bordeaux-Arcachon, les résultats obtenus avec les différents algorithmes sont présentés dans le Tableau 1.1. Pour des raisons de confidentialité les résultats ont été anonymisés par normalisation, la consommation calculée avec l'algorithme SQP dans le Tableau 1.1 est la valeur utilisée pour la normalisation.

CHAPITRE 1. VALIDATION DES MESURES DES COMPTEURS : GÉNÉRATION DES BORNES DE VALIDATION

TABLEAU 1.1 – Impact de l’algorithme sur la précision et le temps de calcul

Algorithme	Énergie calculée	Temps de calcul
Interior-points	Ne converge pas	- s
SQP	1 p.u.	575 s
Active-Set	0,9989 p.u.	6344 s

Les résultats de l’algorithme « SQP » sont retenus pour obtenir la même précision avec un temps de calcul sensiblement inférieur.

Le pas de calcul de l’algorithme est important pour la précision du résultat final et a également un impact sur le temps de calcul. En résolvant le problème formulé sur le temps avec un pas de discrétisation de 2 secondes et l’algorithme « SQP » pour le cas d’étude du TER Bordeaux-Arcachon, les résultats du tableau 1.2 sont obtenus. On observe comment, avec un pas de calcul de 10^{-3} , on obtient pratiquement le même résultat qu’avec des pas de calcul plus petits, qui prennent beaucoup plus de temps de calcul.

TABLEAU 1.2 – Impact du pas de calcul sur la précision et le temps de calcul avec l’algorithme SQP

Pas de calcul	Énergie calculée	Temps de calcul
10^{-2}	Ne converge pas	- s
10^{-3}	1,0027 p.u.	2380 s
10^{-4}	1 p.u.	2575 s
10^{-5}	0,9999 p.u.	3487 s
10^{-8}	0,9990 p.u.	4500 s

Finalement, il reste à déterminer le pas de discrétisation du problème h . Pour évaluer son impact (Tableau 1.3), on va résoudre le problème formulé sur la base du temps avec l’algorithme « SQP » et un pas de calcul de 10^{-3} .

TABLEAU 1.3 – Impact du pas de discrétisation h sur la précision et le temps de calcul avec l’algorithme SQP

Pas de discrétisation h	Énergie calculée	Temps de calcul
10 s	0,7152 p.u.	54 s
6 s	0,8446 p.u.	244 s
5 s	0,8816 p.u.	331 s
3 s	0,9572 p.u.	1146 s
2 s	1 p.u.	2440 s
1 s	1,0411 p.u.	14361 s

L'importance du choix du pas de discrétisation transparaît clairement dans la précision des tests réalisés. De même, une croissance du temps de calcul est aussi observable lorsque diminue le pas de discrétisation. Le temps de calcul peut être réduit sous certaines conditions. La première est de connaître à un instant t la vitesse et la position du train (par exemple les arrêts dans les gares) pour pouvoir appliquer l'optimisation à des parcours plus courts. La deuxième condition est de bénéficier d'un processeur avec plusieurs cœurs dans l'ordinateur où l'on utilise MATLAB. De cette façon, on pourra lancer dans chaque cœur le calcul d'un parcours.

1.4 Maximum «ferroviaire», calcul de la borne de validation supérieure

L'estimation de la consommation maximale atteignable pour un trajet n'est pas analysée dans la littérature. En effet, les études réalisées jusqu'à présent ont pour objectif de rendre les systèmes ferroviaires plus performants. Or nous cherchons dans cette thèse à élaborer des outils mathématiques pour valider et estimer la consommation réelle, performante ou pas. Néanmoins, le même outil mathématique (la « commande optimale ») conçu pour trouver une consommation minimale peut être utilisé pour trouver la consommation maximale. Il suffit de multiplier par -1 la fonction objectif pour trouver la consommation maximale. La consommation maximale ne s'obtient pas en ne conduisant qu'à vitesse maximale, comme on peut l'observer dans la Figure 1.4 pour le cas d'étude du TER Bordeaux-Arcachon. En effet, en conduisant à la vitesse maximale, le trajet s'effectue en un temps inférieur à celui disponible. La consommation maximale pour un temps de trajet donné est obtenue en accélérant et freinant tout le temps pour faire augmenter la consommation (ce qui revient à consommer de l'énergie électrique pour la dissiper sous forme de chaleur dans les freins), si bien que si ce cas est très improbable, il est néanmoins tout à fait réalisable.

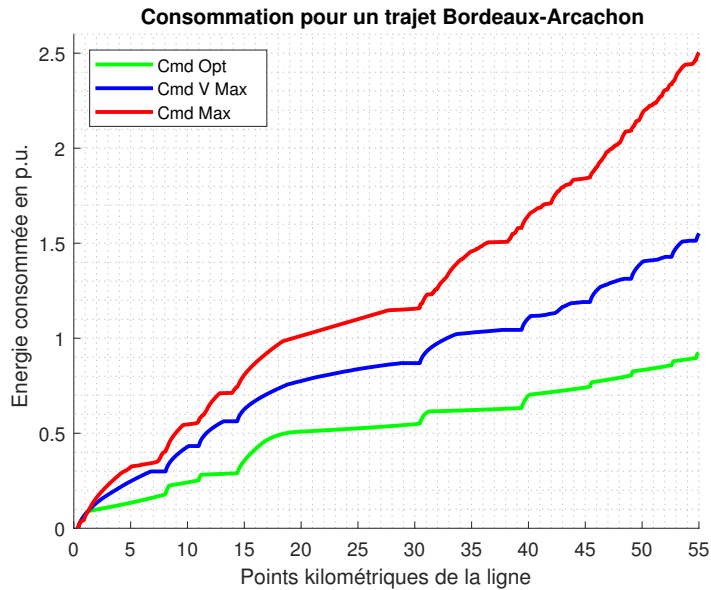


FIGURE 1.4 – Consommation minimale (en vert), consommation maximale (en rouge) et à vitesse maximale (en bleu) pour un trajet donné

Le style de conduite nécessaire pour obtenir la consommation maximale représentée par la courbe rouge dans la Figure 1.4 requiert toutefois beaucoup d’actions sur la commande de conduite en passant presque en continue entre la traction maximale et le freinage maximal comme on peut l’observer dans la Figure 1.5. Donc pour estimer une consommation maximale plus logique d’un point de vue ferroviaire, une autre stratégie a été appliquée. Dans ce type de conduite, le train accélère jusqu’à la vitesse maximale, puis freine jusqu’à une vitesse minimale déterminée (VL) puis accélère à la vitesse maximale, etc. La vitesse minimale VL est définie comme la vitesse minimale à laquelle le train peut développer sa puissance maximale. En restant entre cette vitesse et la vitesse maximale le train peut consommer plus d’énergie en un temps plus court. Comme on peut l’observer dans la Figure 1.6, le train développe sa puissance maximale à 43 km/h, ce sera donc la vitesse retenue comme la vitesse minimale dans notre contrôle.

1.4. MAXIMUM «FERROVIAIRE», CALCUL DE LA BORNE DE VALIDATION SUPÉRIEURE

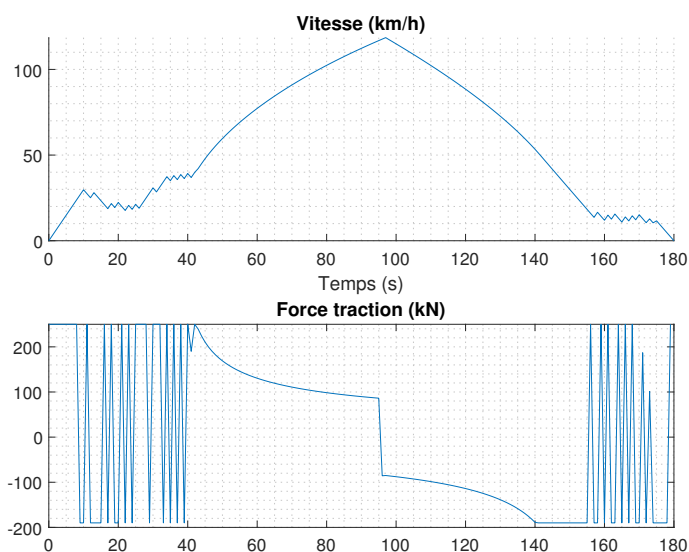


FIGURE 1.5 – Profil de vitesse et commande pour une consommation maximale dans le sens mathématique

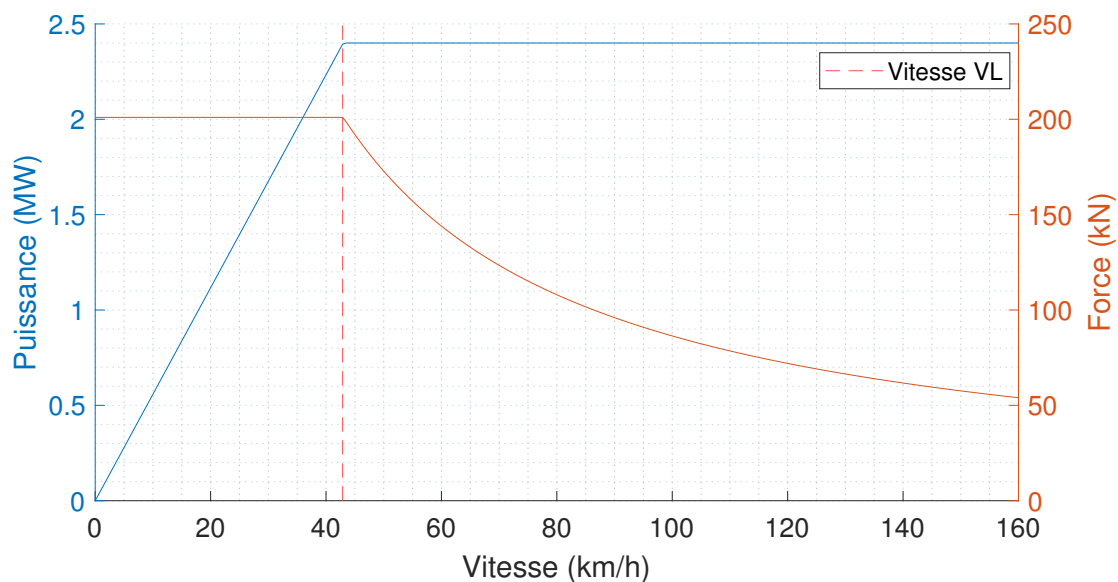


FIGURE 1.6 – Puissance développée par le train dans toute sa plage de vitesse

Dans la Figure 1.7, le profil de vitesse obtenu avec cette nouvelle commande peut être observé. Le train accélère jusqu'à la vitesse maximale pour ensuite freiner jusqu'à la vitesse minimale. Par la suite, le train accélère une nouvelle fois pour atteindre la vitesse maximale, mais il doit commencer à freiner plus tôt pour arriver à une vitesse nulle à l'arrêt suivant.

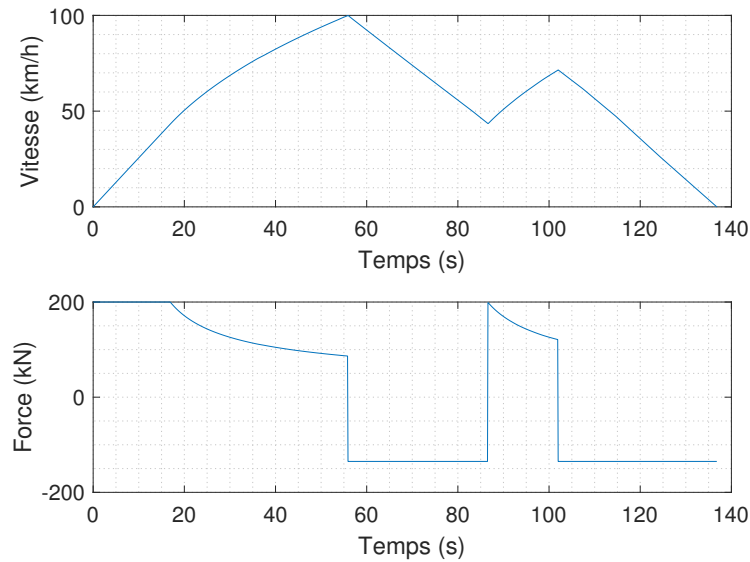


FIGURE 1.7 – Profil de vitesse et de force pour un trajet avec une consommation maximale « ferroviaire »

Le calcul de ce point de freinage n'est pas trivial. Le modèle comprenant plusieurs non-linéarités, une résolution analytique est écartée. En outre, une résolution à travers des approches successives peut devenir très longue à calculer. La technique de l'inversion des modèles peut dès lors représenter la solution, avec deux simulations, ce point de freinage peut être calculé. Le modèle physique du train est facilement réversible, il suffit de remplacer les forces résistives par des forces de traction et la force de traction par une force résistive. En plus, on inversera les limites de traction et de freinage entre elles. Finalement, il ne reste qu'à faire parcourir le train dans le sens inverse. En croisant les deux résultats dans le plan vitesse/espace, on peut facilement trouver le point de freinage. En fait, le point de croisement des deux courbes est le point de freinage recherché. Dans la Figure 1.7.a, la superposition des courbes peut être observée ainsi que la détermination du point de freinage. Dans la Figure 1.7.b, le résultat final est mis en évidence avec l'apport de chaque simulation.

1.4. MAXIMUM «FERROVIAIRE», CALCUL DE LA BORNE DE VALIDATION SUPÉRIEURE

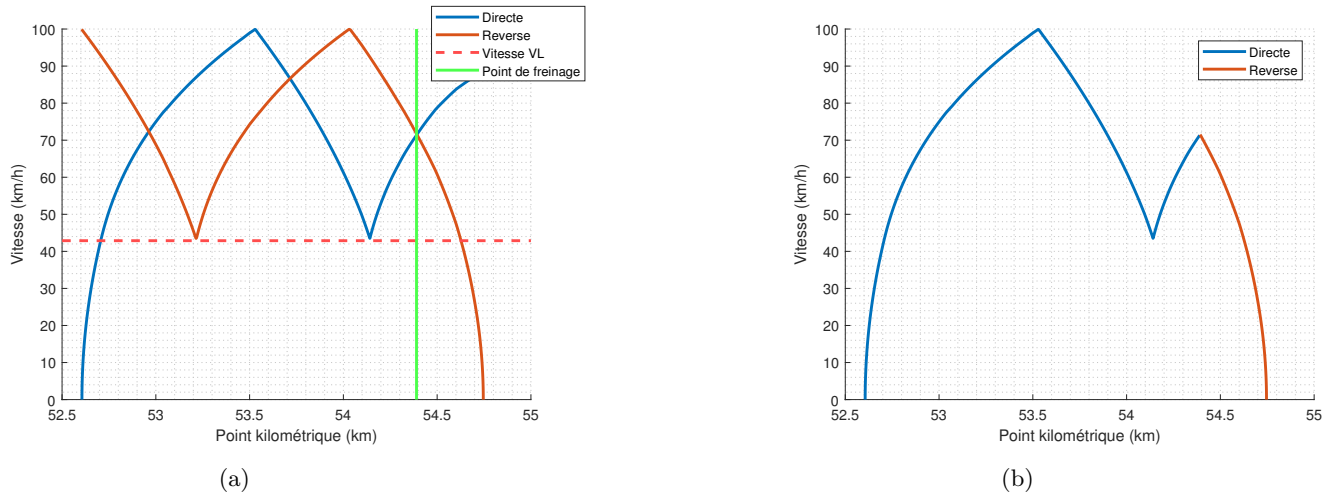


FIGURE 1.7 – Calcul du point de freinage avec l'inversion de modèle

Dans la Figure 1.8 apparaît la nouvelle consommation calculée (en orange). Cette nouvelle consommation est bien supérieure à une consommation à vitesse maximale (en bleu), mais inférieure à une consommation maximale du point de vue mathématique (en rouge). On peut donc dire que la consommation calculée avec la dernière méthode est une consommation maximale réaliste d'un point de vue ferroviaire.

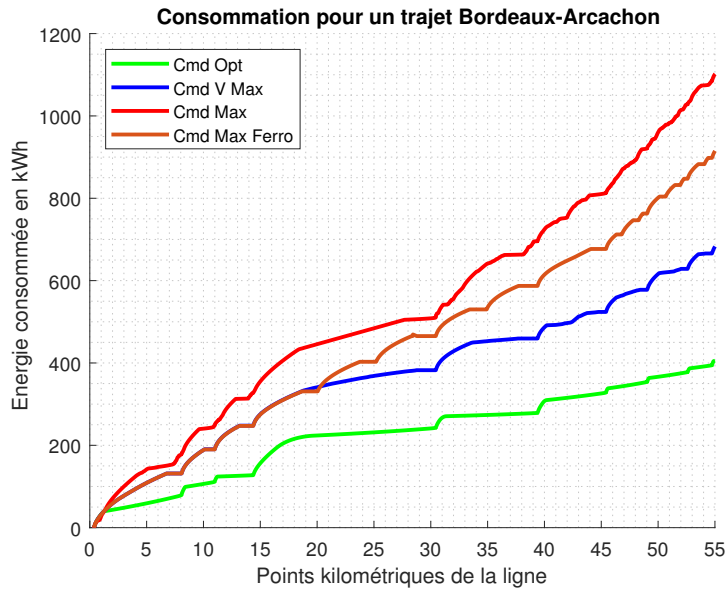


FIGURE 1.8 – Consommation minimale (en vert), à vitesse maximale (en bleu), consommation maximale « ferroviaire » (en orange) et consommation maximale « mathématique » (en rouge) pour un trajet donné

1.5 Calcul des bornes maximales et minimales avec un temps de trajet variable

Le choix d'un temps fixe de parcours n'est pas en accord avec la réalité; les différents trains, même s'ils ont une grille horaire à respecter, accomplissent les trajets en des temps différents. Pour se rapprocher le plus de la réalité, une démarche de calcul des bornes maximales et minimales avec un temps de parcours variable a été développée. Son objectif est d'établir un graphique identique à celui de la Figure 1.1. Le trait rouge de la Figure 1.1 représente la consommation maximale ferroviaire avec la conduite proposée dans la Figure 1.8 en orange et le trait orange de la Figure 1.1 représente la consommation à vitesse maximale de la Figure 1.8 en bleu. Le trait vert continu représente la consommation minimale avec un taux de récupération nul. Grâce à ces graphiques, on peut valider d'un point de vue physique les données de comptage relevées. Pour qu'une consommation soit acceptable il est attendu que les consommations d'énergies mesurées sur le trajet soient positionnées sur le graphe entre le maxima et le minima. A l'inverse, une consommation hors de cet intervalle est significative d'une mesure erronée.

Trouver la consommation minimale pour le temps de parcours minimal est trivial puisque cela ne peut se faire qu'à la vitesse maximale. Par contre, dès que le temps de parcours est rallongé, le problème augmente très rapidement en complexité à cause des plusieurs possibilités qui se créent avec chaque rallongement. En effet, l'étude de la répartition du temps de parcours « extra » qui peut être donné à un train est un sujet complexe largement traité par la littérature [12][27][28]. Différentes approches y sont proposées pour résoudre le problème de répartition du temps. Dans la thèse, l'approche suivante a été retenue :

- Un temps minimal de parcours entre toutes les gares (10 pour le cas d'étude 1) est fixé.
- Le temps de parcours est augmenté par pas de 30 secondes jusqu'à arriver à deux fois le temps nominal de parcours.
- La consommation maximale et minimale que l'on peut avoir pour les différents trajets entre gares et temps de parcours est calculée avec les méthodes utilisées jusqu'à maintenant.

Une fois que la consommation est calculée pour les différents trajets entre gares, il ne reste qu'à les combiner avec les différents temps de parcours pour trouver la combinaison donnant les consommations minimale et maximale et reconstruire la consommation présentée dans la Figure 1.1. Même si dans un premier temps, une résolution par « force brute » (le calcul de toutes les possibilités) semble la solution la plus facile, un calcul de complexité montre qu'un très grand nombre de calculs doivent être réalisés :

1.5. CALCUL DES BORNES MAXIMALES ET MINIMALES AVEC UN TEMPS DE TRAJET VARIABLE

$$combi_{trajet_1} \cdot combi_{trajet_2} \cdot \dots \cdot combi_{trajet_{n-1}} \cdot combi_{trajet_n} = 317.915.136 \text{ combinaisons}$$

Pour résoudre le problème, on devrait privilégier des approches d'optimisation. Le problème peut être modélisé sous Matlab de la façon suivante :

$$\begin{aligned} \min \quad & \sum_{i=1}^n Conso(t_i, n) \\ \text{s.l.c.} \quad & i \in [1, i_{max}(n)] \\ & \sum_{i=1}^n t_i = t_{trajet} \end{aligned} \tag{1.8}$$

Dans Matlab la fonction *patternsearch* a été utilisée pour résoudre le problème (1.8). Cette fonction ne trouvant que des minimums locaux, une procédure « Multi Start » a été rajoutée pour bien couvrir l'espace de recherche. En outre, une comparaison entre le minimum calculé par optimisation et le vrai minimum (calculé par la méthode « force brute ») pour un trajet avec 8 arrêts dans le cas d'étude Bordeaux-Arcachon a été effectuée. Comme nous l'observons dans la Figure 1.9, les deux réponses collent presque parfaitement. En effet, si l'optimisation n'arrive pas à trouver la répartition pour la vraie consommation minimale, l'écart entre le minimum trouvé et le vrai minimum est néanmoins négligeable.

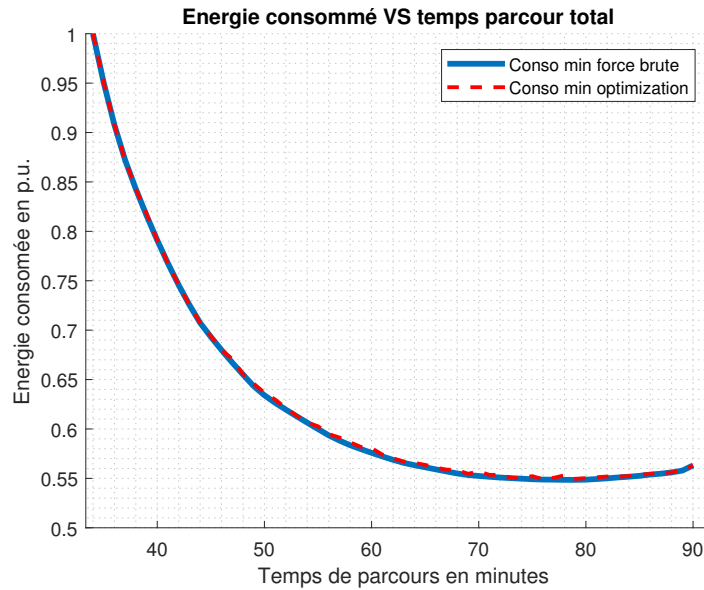


FIGURE 1.9 – Comparaison entre le minimum calculé par la force brute et par optimisation pour un trajet de 8 arrêts

CHAPITRE 1. VALIDATION DES MESURES DES COMPTEURS : GÉNÉRATION DES BORNES DE VALIDATION

Dans la Figure 1.10, les bornes maximales et minimales obtenues peuvent être observées. Pour trouver la limite supérieure de la consommation, nous avons choisi le style de conduite le plus consommateur, avec la masse maximale du train et une consommation maximale des équipements auxiliaires. Pour la consommation minimale, nous avons envisagé le style de conduite le moins consommateur, avec une masse minimale du train et une consommation minimale des auxiliaires. Sur cette figure, les consommations relevées pendant les années 2016 (ronds bleus) et 2018 (croix rouges) sont aussi représentées. Les consommations pour un temps de trajet supérieur à 100 minutes (2 fois le temps du trajet normal) peuvent être causées par des cas exceptionnels (incidents de voie, problèmes dans le train, etc.) ou par des erreurs dans la détermination du trajet dans le traitement des données (liées à l'incertitude d'échantillonnage de 5 min des compteurs). Le cas échéant, ces données ne sont pas valides pour les bornes max/min. D'un autre côté, quelques mesures de 2016 sont très proches du temps minimal de parcours, mais en dessous de la courbe de consommation minimale, ce qui peut s'expliquer par le fait que la rame ne s'est pas arrêtée dans les gares intermédiaires. En effet, en 2016 toutes les rames effectuant ce trajet n'étaient pas livrées. Ces circulations d'autour 40 min seraient donc liées à des circulations de validation du matériel ou à l'habilitation des conducteurs. Dans les années 2017, 2018 et 2019, on ne trouve pas des trajets si courts.

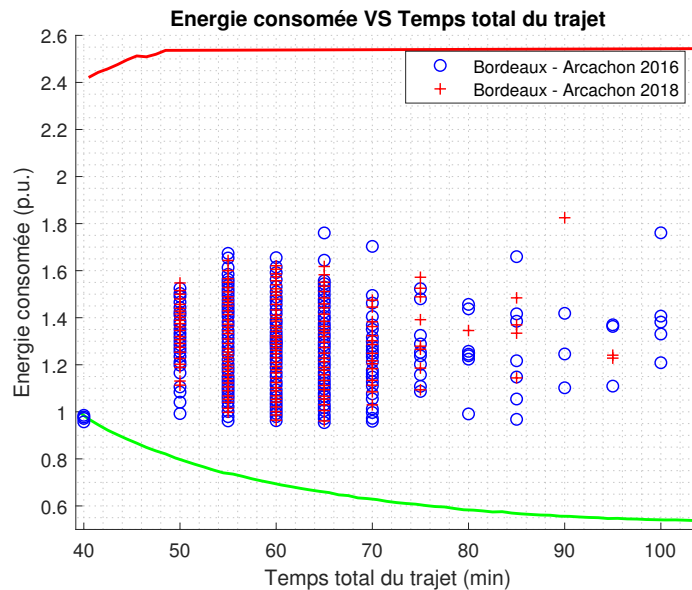


FIGURE 1.10 – Bornes maximales et minimales pour un trajet Bordeaux-Arcachon avec différents temps de parcours

1.6 Application à des cas d'étude

Dans cette section, la méthode développée sera testée dans trois cas d'étude. Nous nous sommes limités à mener l'étude sur les TER parce qu'ils ont une courte durée de trajet entre arrêts et que le problème a de ce fait une taille gérable (temps de trajet entre gares autour de la dizaine de minutes), ce qui n'est pas le cas des trains de type FRET qui ont de longs temps de parcours entre arrêts (plusieurs heures contre une dizaine de minutes pour les TER).

Dans tous les cas d'étude, les consommations des compteurs pour 3 années sont disponibles (2017, 2018 et 2019). Les bornes générées devront en principe valider toutes les mesures enregistrées puisque les compteurs installés sont assez récents et il est attendu qu'ils continuent de mesurer les énergies conformément à leur classe de précision. Toutes les consommations ont une durée de trajet associée, cette durée a une incertitude de ± 5 min au départ et à l'arrivée du fait de la méthode utilisée pour déterminer les trajets. En effet, elle fait appel aux positions GPS données par les compteurs pour déterminer le départ/arrivée du train dans les gares de départ/terminus, mais ces positions GPS n'étant récupérables que toutes les 5 min, des incertitudes de ± 5 min sont générées.

1.6.1 Cas d'étude TER n°1 Bordeaux-Arcachon

Dans ce cas d'étude, les trains parcourent la ligne entre Bordeaux et Arcachon d'une longueur de 58,17 km et comptant 11 gares. La ligne a deux limites de vitesse : 160 km/h pour les premiers 45 km au départ de Bordeaux et 100km/h pour les derniers 13 km. La ligne ne subit pas de grands changements d'altitude : Bordeaux est situé à 7 mètres au-dessus du niveau de la mer et Arcachon à 8 mètres, le point le plus haut de la ligne étant situé à 60 mètres au-dessus du niveau de la mer.

CHAPITRE 1. VALIDATION DES MESURES DES COMPTEURS : GÉNÉRATION DES BORNES DE VALIDATION

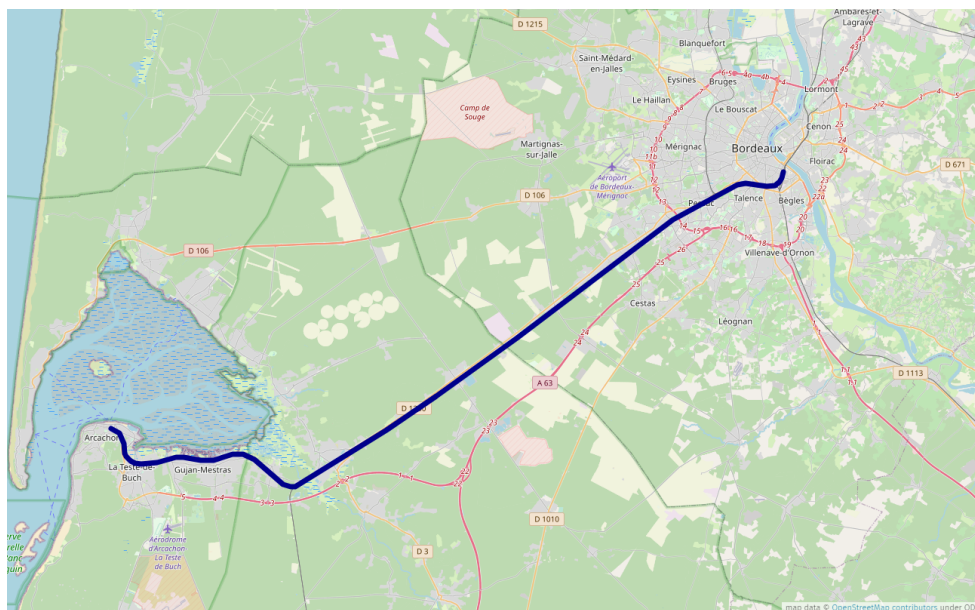


FIGURE 1.11 – Situation géographique de la ligne étudié

Les temps de parcours minimal, nominal (d'après la fiche horaire) et maximal entre arrêts sont indiqués dans le Tableau 1.4.

TABLEAU 1.4 – Temps de parcours entre les arrêts de la ligne Bordeaux-Arcachon

Trajet	Temps minimal	Temps nominal	Temps maximal
Bordeaux - Pessac	5 min	6 min	12 min
Pessac - Allouette	3 min	3 min	6 min
Allouette - Cestas	3 min	4 min	8 min
Cestas - Marcheprime	8 min	9 min	18 min
Marcheprime - Biganos	5 min	8 min	16 min
Biganos - Teich	4 min	5 min	10 min
Teich - Gujan	3 min	5 min	10 min
Gujan - La Hume	3 min	5 min	10 min
La Hume - La Teste	2,5 min	3 min	6 min
La Teste - Arcachon	3 min	5 min	10 min

À l'aide de la méthode développée, les bornes supérieure et inférieure ont pu être déterminées. Dans la Figure 1.12 les bornes inférieures sont montrées pour les deux sens de circulation, et les bornes supérieures sont montrées dans la Figure 1.13. La différence entre les bornes inférieures selon le sens du trajet est en moyenne de 0,52 % avec une différence maximale de 1,74 %. Dans le cas des bornes supérieures, cette différence moyenne est de 4,28 % avec une différence maximale de 5,01 %. Si la différence entre le sens de circulation pour les bornes inférieures est négligeable, ce n'est pas le cas

1.6. APPLICATION À DES CAS D'ÉTUDE

des bornes supérieures. Cela pourrait s'expliquer par la stratégie de conduite choisie pour déterminer la consommation maximale. Pour valider les consommations la borne supérieure la plus grande a été retenue.

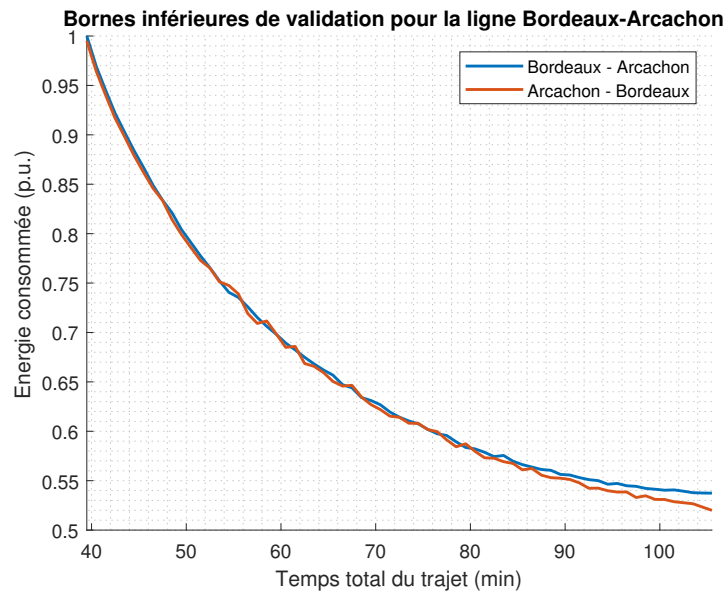


FIGURE 1.12 – Bornes de validation inférieures pour les deux sens de circulation de la ligne Bordeaux-Arcachon

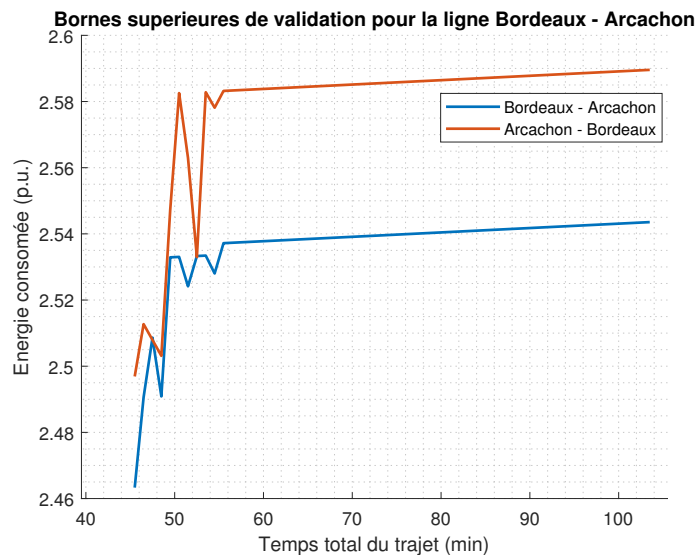


FIGURE 1.13 – Bornes de validation supérieures pour le deux sens de circulation de la ligne Bordeaux-Arcachon

Dans la Figure 1.14 figurent les mesures disponibles pour les années 2017, 2018 et 2019 ainsi que

CHAPITRE 1. VALIDATION DES MESURES DES COMPTEURS : GÉNÉRATION DES BORNES DE VALIDATION

les bornes de validation. Les trajets réalisés dans le sens Bordeaux-Arcachon sont représentés avec des étoiles noires et 100 % de ces mesures sont validées par les bornes. Pour le sens Arcachon-Bordeaux, les consommations sont représentées avec des croix bleues, et 100 % de ces mesures sont validées par les bornes.

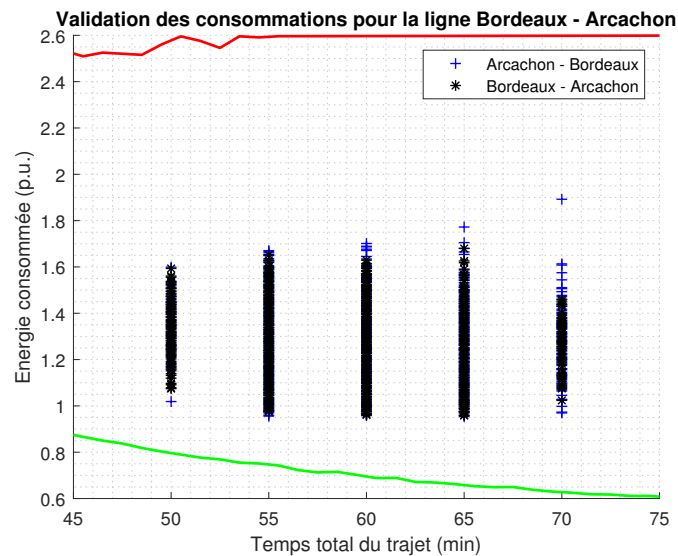


FIGURE 1.14 – Bornes de validation supérieures pour les deux sens de circulation de la ligne Bordeaux-Arcachon avec les mesures des compteurs pour les années 2017, 2018 et 2019

1.6.2 Cas d'étude TER n°2 Bordeaux-Dax

Dans ce cas d'étude, les trains parcourent la ligne entre Bordeaux et Dax, d'une longueur de 147,09 km et comptant 7 gares. La ligne a une seule limite de vitesse, 160 km/h, et ne présente pas de grands changements d'altitude avec Bordeaux à 7 mètres au-dessus du niveau de la mer et Dax à 11 mètres, le point le plus haut de la ligne étant situé à 90 mètres au-dessus du niveau de la mer.



FIGURE 1.15 – Situation géographique de la ligne étudiée

Les temps de parcours minimal, nominal (d'après la fiche horaire) et maximal entre les arrêts figurent dans le Tableau 1.5. Au regard des longs temps de parcours nominal entre les arrêts, le temps maximal n'a pas été calculé comme le double du temps de parcours nominal comme dans le cas d'étude précédent. Une marge de 5 min sur le temps de parcours nominal a été rajoutée pour calculer le temps maximal de parcours. En effet, dans le cas d'étude précédent, le temps de parcours entre arrêts était de l'ordre de quelques minutes, donc doubler le temps de parcours pour le pire cas ne revenait qu'à ajouter 5 minutes ou moins au temps de parcours dans la majorité des cas. Dans le cas d'étude présent, avec des temps de parcours nominaux qui dépassent largement la dizaine voire la vingtaine de minutes, le fait de doubler les temps de parcours pour le pire cas induirait des temps de parcours très longs, éloignés de la réalité et difficiles à calculer.

CHAPITRE 1. VALIDATION DES MESURES DES COMPTEURS : GÉNÉRATION DES BORNES DE VALIDATION

TABLEAU 1.5 – Temps de parcours entre les arrêts de la ligne Bordeaux-Dax

Trajet	Temps minimal	Temps nominal	Temps maximal
Bordeaux - Pessac	4 min	6 min	11 min
Pessac - Biganos	15 min	20 min	25 min
Biganos - Ychoux	12,5 min	18 min	23 min
Ychoux - Labouheyre	6 min	8 min	13 min
Labouheyre - Morcenx	8,5 min	10 min	15 min
Morcenx - Dax	15,5 min	26 min	31 min

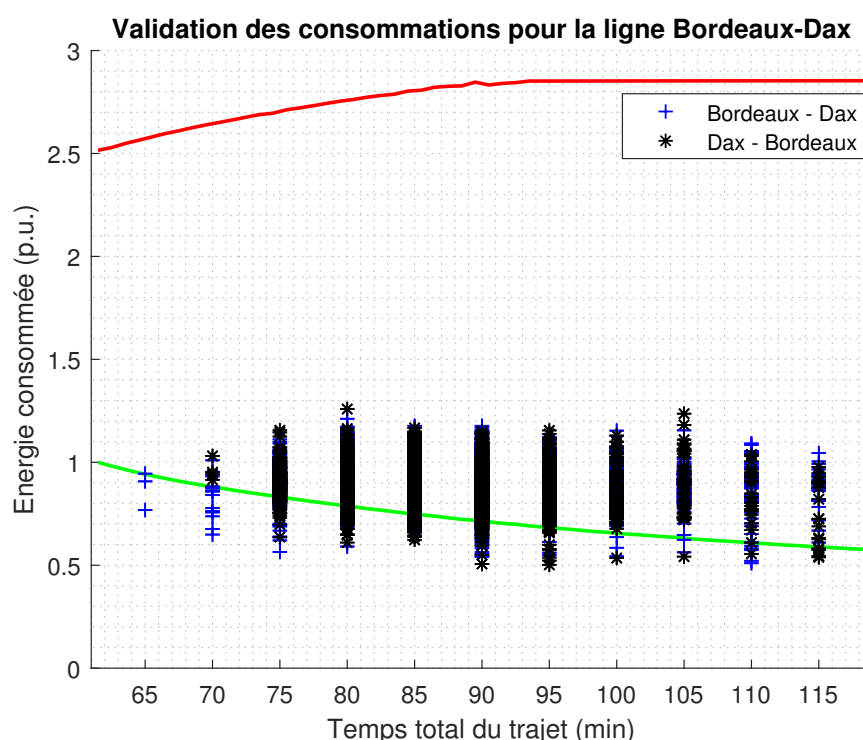


FIGURE 1.16 – Bornes de validation supérieures pour les deux sens de circulation de la ligne Bordeaux-Dax avec les mesures des compteurs pour les années 2017, 2018 et 2019

Dans la Figure 1.16 figurent les mesures disponibles pour les années 2017, 2018 et 2019 ainsi que les bornes de validation. Les trajets réalisés dans le sens Dax-Bordeaux sont représentés avec des étoiles noires et 92,83 % de ces mesures sont validées par les bornes. Pour le sens Bordeaux-Dax, les consommations sont représentées avec des croix bleues, et 96,02 % de ces mesures sont validées par les bornes. Cette fois, le nombre de trajets non validés est plus important que dans les cas précédents. Dans cette ligne TER, deux types de services sont présents, un avec 6 arrêts entre Bordeaux et Dax et un autre avec seulement 4 arrêts. Le fait de ne pas s'arrêter dans toutes les gares rend possible le fait d'obtenir une consommation inférieure à celle d'un service qui s'arrête dans tous les arrêts. A date

d'aujourd'hui on n'est pas capable de différencier ces types de services dans notre base de données. En plus, avec la méthode retenue dans cette thèse, le calcul des bornes de validation pour les services avec 4 arrêts est très difficile dû aux longs temps de parcours entre gares.

Afin d'essayer de valider la supposition que les mesures non validées par la borne inférieure sont des trajets avec 4 arrêts, dans la Figure 1.17 et la Figure 1.18 une tentative de distinction des deux types de services est montrée. Pour faire la différence entre les deux types de service, la grille horaire d'hiver de 2021 est prise. Cette tentative fait la forte hypothèse que les horaires ont été les mêmes pour les années 2017, 2018 et 2019.

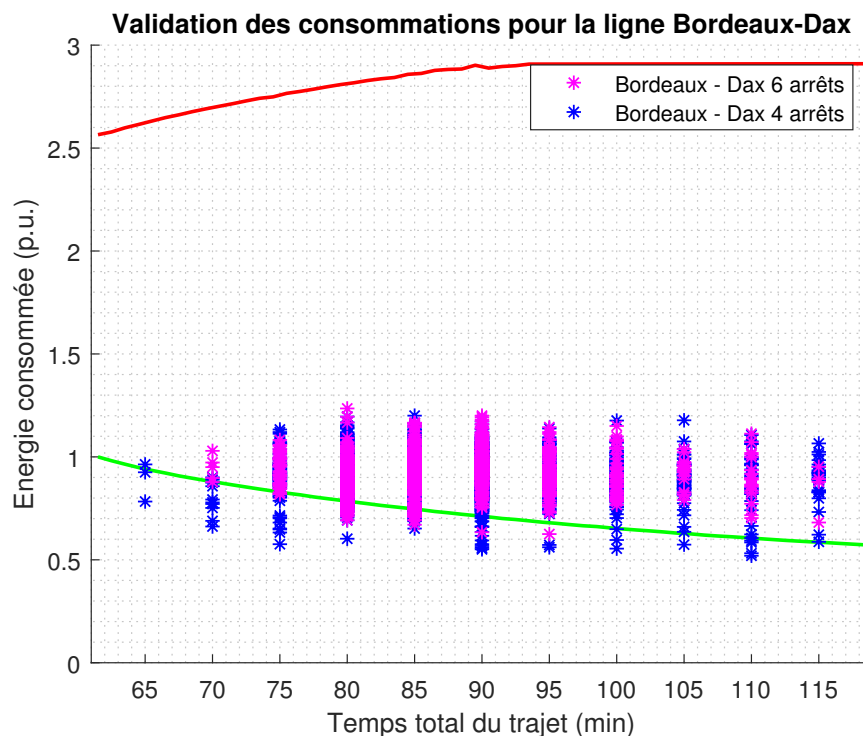


FIGURE 1.17 – Bornes de validation supérieures pour la ligne Bordeaux-Dax avec les mesures des compteurs pour le sens de circulation Bordeaux-Dax pour les années 2017, 2018 et 2019

Dans le sens Bordeaux-Dax un total de 184 trajets ne sont pas validés, 54 trajets (29,35 %) ont été identifiés comme des trajets avec 6 arrêts et 130 trajets (70,65 %) ont été identifiés comme des trajets avec 4 arrêts. Dans le sens contraire un total de 248 trajets ne sont pas validés, 63 trajets (25,40 %) ont été identifiés comme des trajets avec 6 arrêts et 185 trajets (74,60 %) ont été identifiés comme des trajets avec 4 arrêts. En théorie, le 100 % des trajets non validés devraient être des trajets avec 4 arrêts. Cet écart est dû sûrement à la mauvaise identification des différents types de trajets, en effet l'hypothèse sur un horaire constant entre 2017 et 2021 est très forte, sachant que les horaires évoluent naturellement avec les saisons, que la ligne a été impactée par les grèves de 2018 et 2019 ainsi que par

des travaux de rénovation de voies et de la signalisation qui ont impacté le trafic.

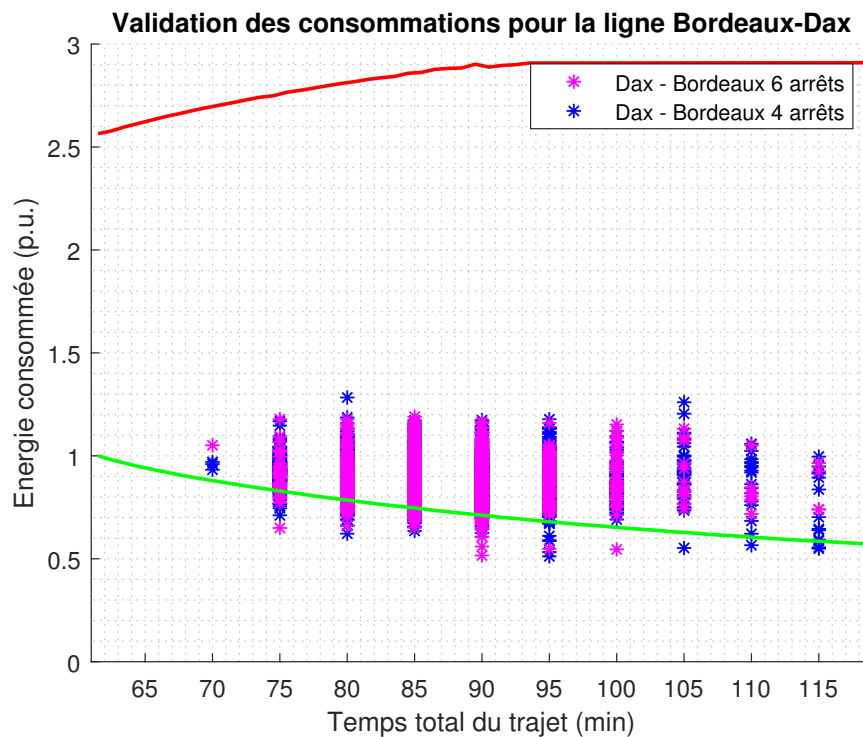


FIGURE 1.18 – Bornes de validation supérieures pour la ligne Bordeaux-Dax avec les mesures des compteurs pour le sens de circulation Dax-Bordeaux les années 2017, 2018 et 2019

1.6.3 Cas d'étude TER n°3 Dijon-Besançon

Dans ce cas d'étude, les trains parcourent la ligne entre Dijon et Besançon, d'une longueur de 91,12 km et comptant 10 gares. La ligne a une seule limite de vitesse : 160 km/h et deux types d'électrification : en courant continu entre Dijon et Dole (45,88 km) et en courant alternatif entre Dole et Besançon (45,24 km). La ligne ne présente pas de grands changements d'altitude avec Dijon culminant à 246 mètres au-dessus du niveau de la mer et Besançon à 283 mètres.

1.6. APPLICATION À DES CAS D'ÉTUDE

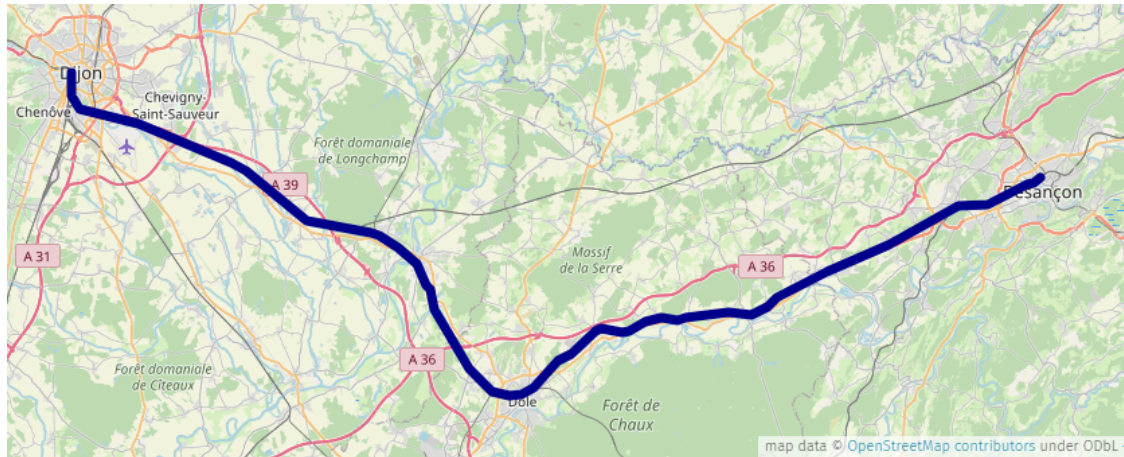


FIGURE 1.19 – Situation géographique de la ligne étudiée

Les temps de parcours minimal, nominal (d'après la fiche horaire) et maximal entre les arrêts est indiqué dans le Tableau 1.6.

TABLEAU 1.6 – Temps de parcours entre les arrêts de la ligne Dijon-Besançon

Trajet	Temps minimal	Temps nominal	Temps maximal
Dijon - Genlis	8 min	11 min	22 min
Genlis - Collonges	2,5 min	4 min	8 min
Collonges - Auxonne	4,5 min	7 min	14 min
Auxonne - Dole	6,5 min	10 min	20 min
Dole - Orchamps	6,5 min	10 min	20 min
Orchamps - Ranchot	3 min	4 min	8 min
Ranchot - Saint-Vit	3,5 min	6 min	12 min
Saint-Vit - Dannemarie	3 min	5 min	10 min
Dannemarie - Besançon	5,5 min	9 min	18 min

A l'aide de la méthode développée, les bornes supérieure et inférieure ont pu être déterminées. Dans la Figure 1.20, les bornes inférieures sont indiquées pour les deux sens de circulation et la Figure 1.21 présente les bornes supérieures. La différence entre les bornes inférieures selon le sens du trajet est en moyenne de 0.23 % avec une différence maximale de 1.01 %. Dans le cas des bornes supérieures, cette différence moyenne est de 5.99 % avec une différence maximale du 7.38 %. À nouveau, si la différence entre le sens de circulation pour les bornes inférieures est négligeable, ce n'est pas le cas pour les bornes supérieures. Cela pourrait s'expliquer par la stratégie de conduite choisie pour déterminer la consommation maximale et sa forte non-linéarité. Pour valider les consommations, c'est la borne supérieure la plus grande qui a été retenue.

CHAPITRE 1. VALIDATION DES MESURES DES COMPTEURS : GÉNÉRATION DES BORNES DE VALIDATION

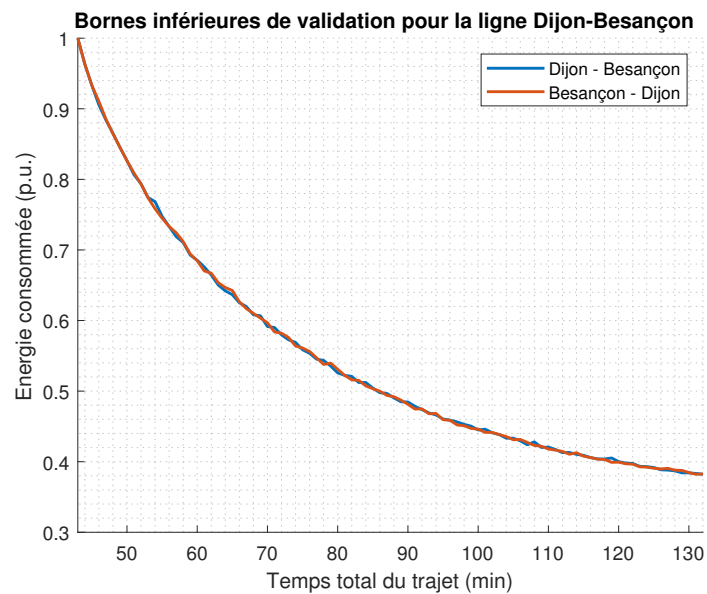


FIGURE 1.20 – Bornes de validation inférieures pour les deux sens de circulation de la ligne Dijon-Besançon

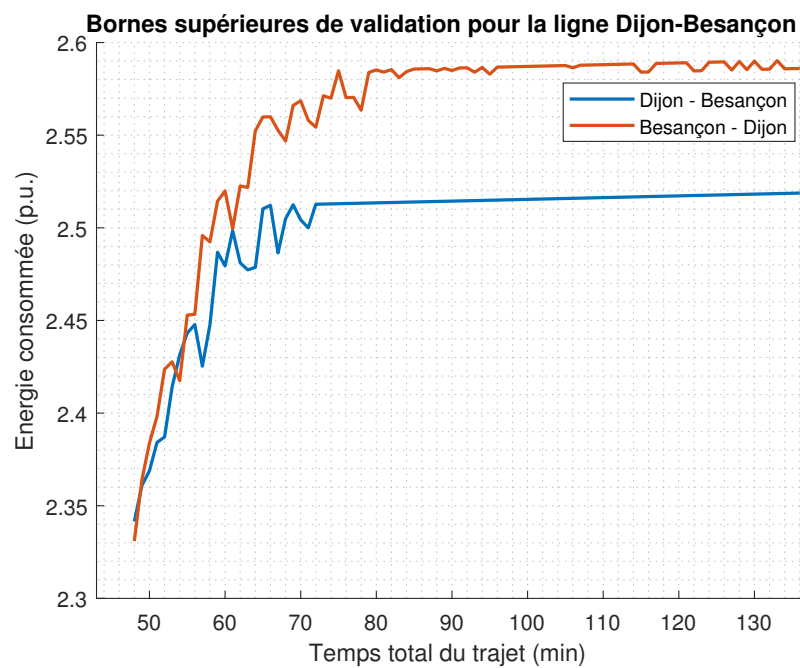


FIGURE 1.21 – Bornes de validation supérieures pour les deux sens de circulation de la ligne Dijon-Besançon

La Figure 1.22 montre les mesures disponibles pour les années 2017, 2018 et 2019 ainsi que les bornes de validation. Les trajets réalisés dans le sens Dijon-Besançon sont représentés par des étoiles noires et 99,93 % de ces mesures sont validées par les bornes. Pour le sens Besançon-Dijon, les consommations sont représentées par des croix bleues et 99,55 % de ces mesures sont validées par les bornes. Les trajets non validés sont soit liés aux trains « semi-directs » qui ne marquent pas tous les arrêts, ils ne s'arrêtent que sur 6 des 10 arrêts de la ligne. Comme indiqué précédemment, le fait de ne pas s'arrêter dans toutes les gares peut occasionner une consommation inférieure à celle qu'on a calculée. A nouveau, on n'est pas capable de différencier ces types de services dans notre base de données. La méthode retenue dans cette thèse ne permet pas non plus le calcul des bornes de validation pour les services avec 6 arrêts est très difficile dû aux longs temps de parcours entre gares.

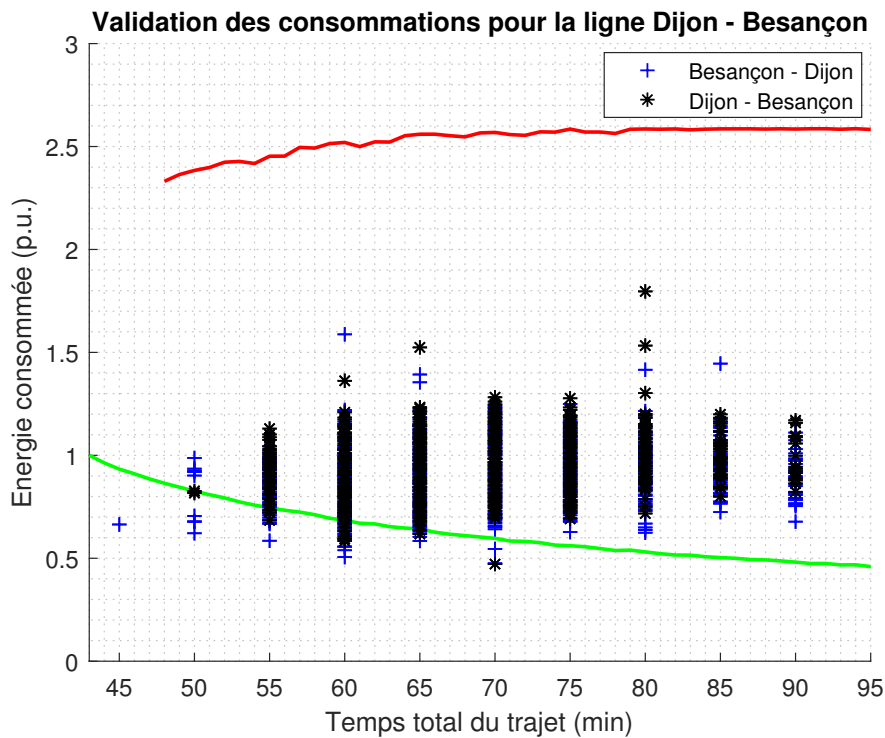


FIGURE 1.22 – Bornes de validation supérieures pour les deux sens de circulation de la ligne Dijon-Besançon avec les mesures des compteurs pour les années 2017, 2018 et 2019

1.7 Conclusion

Dans ce chapitre, une méthode pour générer des bornes de validation à l'aide des modèles physiques pour les consommations enregistrées par les compteurs a été élaborée. Pour la borne inférieure, la modélisation dite « commande optimale » a été choisie pour résoudre le problème. Pour la borne supérieure, une stratégie de conduite plus consommatrice qu'une conduite à vitesse maximale a été développée. Même si cette consommation est inférieure à la consommation maximale du point de vue mathématique, elle est plus proche de la réalité. La reconstruction des bornes pour différents temps de trajet a été aussi traitée. La méthode de validation des consommations utilisée a été validée avec trois cas d'étude. Même si la méthode développée fonctionne, pour les lignes où des trains assurent différents types de service (et donc différents nombre d'arrêts), la création de différentes bornes de validation suivant le type de service assuré par le train peut être justifiée. Le sens de parcours est peu influent dans la détermination des bornes inférieures, mais il l'est dans la détermination des bornes supérieures. Cela est lié à la stratégie de conduite retenue et sa non-linéarité.

Finalement, la méthode développée n'est exploitable que pour des temps de parcours entre arrêts ne dépassant pas une dizaine de minutes. Au-delà, le nombre de variables est trop grand et cet algorithme d'optimisation a de grandes difficultés à résoudre le problème. Cette méthode n'est donc pas appropriée pour la validation des consommations des trains FRET.

Dans le prochain chapitre, des estimateurs de la consommation des trains pour améliorer la disponibilité des données de comptage sont présentés. Ces estimateurs partent d'une base de données composée par des mesures des compteurs validées par la méthode expliquée dans ce chapitre.

Chapitre 2

Estimateurs de la consommation pour un trajet donné

Contenu

1.1	Introduction	26
1.2	Modèles physiques des trains dans la littérature	27
1.3	Contrôle optimal, calcul de la borne de validation inférieure	30
1.3.1	Modélisation mathématique du problème	30
1.3.2	Résolution numérique du problème de « commande optimale » avec MATLAB	31
1.4	Maximum «ferroviaire», calcul de la borne de validation supérieure	35
1.5	Calcul des bornes maximales et minimales avec un temps de trajet variable	40
1.6	Application à des cas d'étude	43
1.6.1	Cas d'étude TER n°1 Bordeaux-Arcachon	43
1.6.2	Cas d'étude TER n°2 Bordeaux-Dax	46
1.6.3	Cas d'étude TER n°3 Dijon-Besançon	50
1.7	Conclusion	54

2.1 Introduction

Dans le chapitre précédent, une méthode qui répondait au premier objectif de cette thèse a été proposée, basée sur des modèles physique et validant le caractère réaliste des consommations enregistrées dans les compteurs. Le deuxième objectif de cette thèse est d'estimer la consommation d'un train pour une journée donnée afin de compenser les possibles erreurs de mesure des compteurs. Une première approche serait de faire appel aux modèles physiques utilisés dans le chapitre précédent pour développer les estimateurs de la consommation. Néanmoins, l'incertitude sur plusieurs paramètres du modèle physique fait que leur utilisation n'est pas adaptée. Les sources des incertitudes (paramètres d'entrée) du modèle physique sont les suivantes : la masse du train, la météorologie (pression atmosphérique et vent), la consommation des auxiliaires (équipements de climatisation principalement), le style de conduite et la « réversibilité » des sous-stations.

Si la répartition des incertitudes des paramètres d'entrée est connue grâce à diverses techniques (la méthode Monte-Carlo notamment), une répartition de l'incertitude à la sortie pourrait être calculée. Néanmoins, dans cette thèse, cette répartition de l'incertitude dans les paramètres d'entrée est inconnue. En revanche, quelques données d'entrée peuvent être éliminées. Par exemple, le terme « réversibilité » des sous-stations peut être éliminé si le calcul se centre strictement sur l'énergie consommée. En disposant des compteurs à deux registres, l'énergie consommée et l'énergie refoulée vers la caténaire, la validation du premier registre suffirait pour valider le bon fonctionnement du compteur. D'un autre côté, une relation de proportionnalité directe ou inverse peut être trouvée entre toutes les entrées restantes et la sortie, permettant ainsi d'identifier le domaine d'incertitude de la sortie du modèle physique, même si la distribution de l'incertitude dans ce domaine ne sera pas connue. En effet, ce domaine est celui défini par les bornes de validation des consommations.

Dans la littérature, l'impact de certains paramètres d'entrée a déjà été évalué. Par exemple, pour l'impact du style de conduite, une variation de $\pm 15\%$ de la consommation totale pour le même trajet pour différents conducteurs est mentionné en [29]. Concernant la masse totale du train et le vent, en prenant l'exemple du cas d'étude TER n°1 Bordeaux-Arcachon, on trouve les écarts montrés dans les Tableau 2.1 et Tableau 2.2 respectivement.

TABLEAU 2.1 – Incertitude du modèle physique liée à la masse totale du train

Masse totale du train	Consommation minimale (p.u)	Consommation maximale (p.u)
Masse minimale, 0 personne/m ² (-24%)	0,8	0,9576
Masse standard, 4 personnes/m ²	1	1
Masse maximale, 6 personnes/m ² (+10%)	1,095	1,015

TABLEAU 2.2 – Incertitude du modèle physique liée au vent

Masse totale du train	Consommation minimale (p.u)	Consommation maximale (p.u)
Orage fort 60 km/h	1,1867	1,0615
Orage léger 30 km/h	1,0767	1,0288
Sans vent 0 km/h	1	1
Orage léger -30 km/h	0,947	0,9778
Orage fort -60 km/h	0,9174	0,9635

2.2 Estimateurs « data-based »

Au regard de la grande incertitude qui ressort du modèle physique, un autre type d'estimateur doit être développé pour estimer la consommation d'un train donné pour un trajet et un jour spécifique. Dans cette thèse, les estimateurs dits « data-based » ont été retenus parce qu'ils ne nécessitent pas de modèle pour obtenir une estimation. Le modèle est directement élaboré à partir d'une base de données, laquelle se base sur les mesures des compteurs validées par la méthode du chapitre précédent.

Dans tous les estimateurs la méthode « cross-validation » sera appliquée. Elle permet d'utiliser tous les points disponibles dans une base de données sans « tricher » (utilisation d'une donnée à la fois pour générer un estimateur et pour en mesurer la performance). Dans cette méthode, les données disponibles pour entraîner l'estimateur sont réparties en « K » groupes. K-1 groupes de données sont utilisés pour générer un estimateur puis la performance de cet estimateur sera testée sur le groupe de données non utilisé pour le générer. Cette procédure de génération des estimateurs est répétée K fois au total, en changeant à chaque fois le groupe sur lequel la performance est calculée. Finalement, pour obtenir la performance de l'estimateur testé, il suffit de faire la moyenne des performances des K estimateurs générés. Dans la thèse, l'indicateur de performance est l'écart-type moyen de l'erreur sur 100 estimateurs générés, ainsi que l'écart-type des écarts-types de ces 100 estimateurs pour mesurer la dispersion des performances individuelles. La valeur de « K » est fixée à 10 pour tous les estimateurs. Ainsi, pour éviter d'introduire des biais dans les estimateurs, une normalisation aléatoire de la quantité des données par mois est faite (les données enlevées des mois surreprésentés ne sont pas toujours les mêmes). Cette normalisation de la base de données est effectuée chaque fois qu'un nouvel estimateur est calculé, donc pour 100 estimateurs générés, 100 processus de normalisation différents seront effectués.

2.2.1 Moyenne sur l'ensemble de la base de données

Cet estimateur est le plus simple à construire. Il sera le référent utilisé pour évaluer les gains dans les performances des autres estimateurs. En partant de la base de données complète (qui peut comprendre les données de plusieurs années) et en utilisant la méthode de «cross-validation», l'estimateur est facilement calculé.

2.2.2 Moyenne mensuelle

La moyenne mensuelle est un estimateur plus complexe qu'une moyenne normale parce qu'elle intègre l'information provenant des mois. Dans cet estimateur, une moyenne mensuelle est calculée, c'est-à-dire que l'on obtient un total de 12 valeurs moyennes pour la consommation. La performance de cet estimateur est calculée en moyennant les performances des 12 estimateurs mensuels.

2.2.3 Régression linéaire multiple (RLM)

La régression linéaire multiple est l'estimateur le plus simple qui permet de prendre en compte plusieurs sources d'information pour estimer plus précisément la consommation des trains. Entre autres sources d'information, on trouve des données calendaires, des horaires, des données météorologiques et des consommations des sous-stations de traction. Cet estimateur est utilisé dans la littérature pour estimer la consommation dans un cas général quand un compteur tombe en panne [30] où dans de cas plus proches au domaine ferroviaire, comme dans [31], où ils estiment la consommation d'un tramway avec des RLM.

2.2.4 Support Vector Machine (SVM)

L'estimateur SVM est l'un des modèles d'intelligence artificielle les plus simples. En comparaison avec une régression linéaire, les SVM permettent d'ajouter des non-linéarités qui aident à mieux s'ajuster aux données d'entraînement. L'estimateur SVM utilise les mêmes données d'entrée que l'estimateur basé sur des régression linéaires multiples. Cet estimateur est utilisé dans [32] pour prédire la consommation du secteur industriel en Afrique du Sud, ou dans [33] où il est utilisé pour détecter la fraude au sein des smart-grids en prédisant les consommations des différents utilisateurs et en comparant avec la mesure communiquée.

2.2.5 Réseaux de neurones

Les réseaux de neurones sont l'estimateur le plus complexe analysé dans cette thèse. Ce type d'estimateur a été utilisé par exemple dans le domaine automobile pour estimer la consommation des moteurs diesel [34] et commence à être utilisé dans le domaine ferroviaire aussi, mais sur une échelle plus macroscopique [35]. Cet estimateur a plusieurs variables à régler : le nombre de couches, le nombre de neurones par couche ainsi que la fonction d'activation des neurones. Après plusieurs essais, la fonction d'activation des neurones du type sigmoïde a été choisie pour sa performance. Dans la Figure 2.1 on montre la topologie retenue pour cette thèse. Le nombre optimal de neurones de la seule couche cachée a été déterminé d'une façon empirique pour chaque cas d'étude.

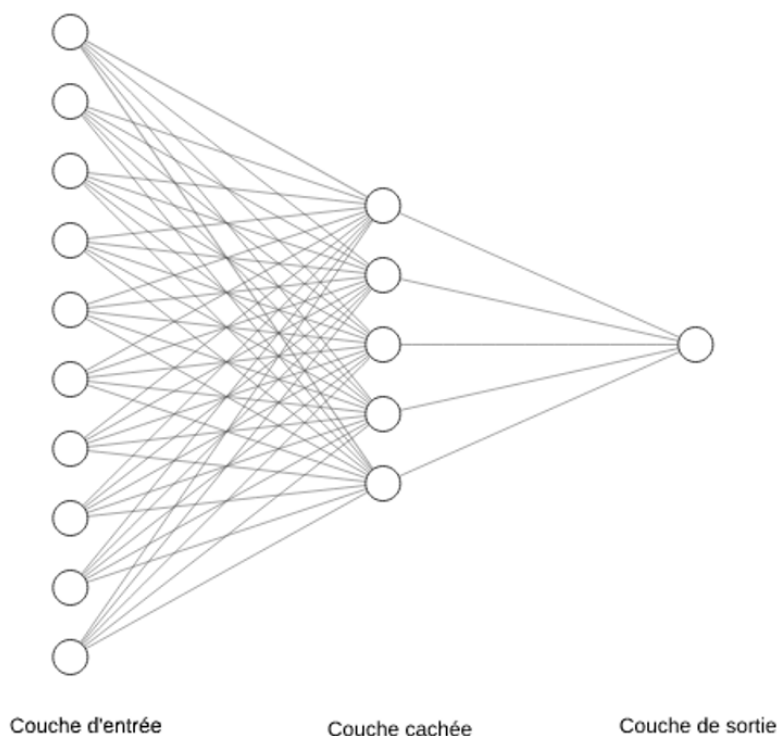


FIGURE 2.1 – Topologie de reseau de neurones utilisée dans cette thèse [36]

2.3 Méthode pour trier les sources d'information disponibles

Toutes les sources d'information disponibles ne sont pas forcément des vraies sources d'information parce que la corrélation entre elles et les consommations peut être nulle. La méthode de tri développée permet d'identifier ces sources de bruit ainsi que l'apport d'information de chaque source, et donc, les sources d'information principales. Cette méthode sera appliquée pour trier les sources d'information pour les estimateurs du type régression linéaire multiple, Support Vector Machine et réseaux de

neurones. Elle fonctionne de la façon suivante : en supposant N sources d'information, N estimateurs sont calculés avec $N-1$ sources d'information (en excluant une source d'information distincte à chaque fois). La source d'information éliminée dans l'estimateur qui a la meilleure performance est identifiée comme la source d'information la moins importante. Si la performance de l'estimateur sans cette source d'information est meilleure qu'avec elle, la source d'information sera identifiée comme une source de bruit. Ensuite, $N-2$ estimateurs sont calculés avec $N-2$ sources d'information (en excluant la source d'information la moins importante déjà identifiée, plus une autre à chaque fois) et l'estimateur avec la meilleur performance identifiée. La source d'information éliminée dans cet estimateur est la deuxième source d'information la moins importante. Ce processus est répété jusqu'à avoir trié toutes les sources d'information.

2.4 Cas d'étude TER

Pour tous les cas d'étude analysés, l'information de la consommation ainsi que les données d'entrée des estimateurs pour 3 ans (2017, 2018 et 2019) ont été récupérées. Dans chaque cas d'étude, deux estimateurs ont été développés, un pour chaque sens de circulation de la ligne.

2.4.1 Sources d'information disponibles

Pour les cas d'étude TER, les sources d'information suivantes sont disponibles :

- Jour de l'année (1 à 365)
- Jour de la semaine (1 à 7)
- Mois (1 à 12)
- Année (2017, 2018 ou 2019)
- WE ou férié (binaire)
- Heure de départ (par troncature)
- Durée du trajet (minutes, précision ± 5 min)
- Température ($^{\circ}\text{C}$)
- Consommation dans les sous-stations lors du passage du train

La durée du trajet a une incertitude de ± 5 min au départ et à l'arrivée du fait de la méthode utilisée pour déterminer les trajets. En effet, elle fait appel aux positions GPS données par les compteurs pour déterminer le départ/arrivée du train dans les gares de départ/terminus, mais ces positions GPS n'étant récupérables que toutes les 5 min, des incertitudes de ± 5 min sont générées.

Pour les températures ambiantes, elles sont récupérées dans l'aéroport le plus proche de la ligne étudiée : l'aéroport de Bordeaux à Mérignac pour les cas d'étude Bordeaux-Arcachon et Bordeaux-Dax et l'aéroport de Dijon-Bourgogne pour le cas d'étude Dijon-Besançon.

Dans les cas des consommations des sous-stations, on adopte le processus suivant pour construire cette source d'information : tout d'abord, il a fallu établir un historique de consommation avec un pas d'échantillonnage de 5 min. Actuellement, les compteurs des sous-stations relèvent la consommation toutes les 10 min. Pour revenir à un pas d'échantillonnage de 5 min, une consommation constante pendant ces 10 min est supposée et donc les pas de 5 min ont été reconstruits en divisant par deux les valeurs de 10 min. Pour calculer la consommation des sous-stations lors du passage du train, il faut distinguer les cas AC et DC. Dans les cas où l'alimentation se fait en AC, le train ne peut consommer que depuis une seule sous-station en même temps parce que les sous-stations sont isolées entre elles, donc des zones d'alimentation pour chaque sous-station peuvent facilement être établies. Avec cette information et la position GPS du train, la consommation dans les sous-stations lors du passage du train est aisément déterminée. Pour les cas où l'alimentation se fait en DC, les sous-stations ne sont pas isolées entre elles et donc un train est normalement alimenté par au moins 2 sous-stations. Outre, en plus de compliquer le processus de reconstruction de l'information de la consommation dans les sous-stations lors du passage du train, cela ajoute du bruit sur ces mesures parce que l'information de la consommation du train est répartie dans plusieurs entrées.

2.4.2 Cas d'étude TER n°1 Bordeaux-Arcachon

Dans ce cas d'étude, 4 sous-stations en DC alimentent la ligne : La Médoquine, Croix, La Mothe et La Teste. Les données disponibles, qui s'étalent sur trois années, établissent un total de 12901 trajets Bordeaux-Arcachon et 13034 trajets Arcachon-Bordeaux. Dans le Tableau 2.3, les performances des estimateurs sont montrées, le pourcentage affiché est l'écart-type moyen de l'erreur de 100 estimateurs et le pourcentage affiché entre parenthèses est l'écart-type de l'écart-type de ces 100 estimateurs. On devrait observer une amélioration des performances des estimateurs avec la montée en complexité de ceux-ci. Cela est vérifié sauf pour le dernier estimateur, le réseau de neurones. Cela peut être dû à la complexité de la mise au point de cet estimateur. Après l'essai avec plusieurs couches et différents nombres de neurones par couche, la configuration optimale trouvée pour ce cas d'étude est une seule couche avec 24 neurones. En regardant l'écart-type de l'écart-type sur les 100 estimateurs, on observe qu'il augmente avec la complexité des estimateurs : plus l'estimateur est complexe, plus il est dépendant de la base de données qu'on lui fournit. Il convient donc de s'assurer que la base de données fournie est bien représentative du modèle et qu'elle n'introduit pas de biais. Dès lors, l'écart-type de l'écart-type sur les 100 estimateurs doit avoir une valeur basse en comparaison avec l'écart-type de l'erreur trouvé.

CHAPITRE 2. ESTIMATEURS DE LA CONSOMMATION POUR UN TRAJET DONNÉ

TABLEAU 2.3 – Performance des estimateurs de la consommation pour le cas d’étude Bordeaux-Arcachon

écart-type moyen de l’erreur (écart-type de l’écart-type sur 100 estimateurs)	Bordeaux-Arcachon	Arcachon-Bordeaux
Moyenne sur l’ensemble de la base de données	10,91% (0,001%)	11,66% (0,001%)
Moyenne mensuelle	10,65% (0,003%)	11,46% (0,004%)
Régression linéaire multiple	10,10% (0,036%)	11,05% (0,037%)
SVM	9,36% (0,048%)	9,96% (0,048%)
Réseaux de neurones	9,43% (0,052%)	9,99% (0,047%)

Un autre point important pour déterminer la viabilité des estimateurs est de calculer le temps de calcul que demande chacun d’entre eux. Cela est analysé de la Tableau 2.4, où le temps de calcul nécessaire pour générer chaque estimateur est montré. La moyenne sur l’ensemble de la base de données étant l’estimateur le plus rapide à calculer, elle a été choisie comme référence. La moyenne mensuelle demande un peu plus de temps de calcul liée à la classification des données par mois. Pour la RLM, le premier estimateur à intégrer des sources d’information, le temps de calcul se double par rapport à la référence. Dans le cas du SVM cette augmentation du temps de calcul est beaucoup plus grand, mais reste toujours faible en comparaison avec le temps de calcul des réseaux de neurones. Ceux-ci nécessitent un temps de calcul élevé à cause des nombreux paramètres qui doivent être fixés. A titre d’exemple, la valeur de l’unité de temps de calcul est de 0.26 secondes dans un processeur Intel i7-8700K.

TABLEAU 2.4 – Temps de calcul pour la generation chaque estimateur pour le cas d’étude Bordeaux-Arcachon

Type de estimateur	Temps de calcul
Moyenne sur l’ensemble de la base de données	1 unité
Moyenne mensuelle	1.3 unités
Régression linéaire multiple	2 unités
SVM	231 unités
Réseaux de neurones	5612 unités

L’apport de chaque source d’information dans l’estimateur SVM a été analysé dans le Tableau 2.5. Ces résultats ont été obtenus grâce à la méthode de tri des sources d’information expliquée précédemment. Les sources d’information qui ont un apport négatif sont identifiées comme des sources de bruit (parce que la performance de l’estimateur est améliorée sans elles) et celles qui ont un apport positif sont identifiées comme de vraies sources d’information. Le pourcentage affiché est le gain/perte

dans l'écart-type moyen de l'erreur de l'estimateur.

Il ressort en premier lieu que la température ambiante est la source d'information la plus importante, ce qui peut être expliqué par la grande consommation des équipements de climatisation du nouveau matériel roulant. L'année de la consommation et l'heure de départ du train sont les deux autres sources d'information les plus importantes. Le premier facteur semble indiquer une évolution de la consommation en lien avec une évolution du nombre de passagers transportées, tandis que le deuxième permet d'identifier les périodes creuses et de pointe.

TABLEAU 2.5 – Apport de chaque source d'information pour l'estimateur SVM pour le cas d'étude Bordeaux-Arcachon

Source d'information	Bordeaux-Arcachon	Arcachon-Bordeaux
Jour année	14,38 %	1,75 %
Jour semaine	Bruit	Bruit
Mois	2,02 %	6,39 %
Année	20,58 %	12,05 %
WE/Férié	Bruit	9,08 %
Heure départ	19,63 %	9,99 %
Durée trajet	6,48 %	Bruit
Température	32,46 %	51,97 %
STA 1 - DC	Bruit	Bruit
STA 2 - DC	Bruit	Bruit
STA 3 - DC	Bruit	Bruit
STA 4 - DC	4,44 %	8,77 %

La consommation récupérée dans les sous-stations lors du passage du train, n'est quant à elle pas identifiée comme une source d'information sauf pour le cas de la sous-station 4. Cette sous-station n'alimente en effet qu'un petit morceau de la ligne qui est exclusivement parcourue par des TER, ce qui laisse entendre que dans les cas des autres sous-stations, les mesures obtenues sont tellement polluées par les consommations des autres trains que la signature de la consommation des TER est perdue.

Pour l'année 2018, les consommations dans les sous-stations avec un pas d'échantillonnage de 1 min sont disponibles. Afin de vérifier si avec cet échantillonnage plus fin, les signatures des TER peuvent être récupérées, de nouvelles données d'entrée de consommations de sous-stations ont été créées. Pour l'année 2018, on compte 4106 trajets Bordeaux-Arcachon et 4152 trajets Arcachon-Bordeaux. Dans le Tableau 2.6 sont mises en évidence les performances des estimateurs avec cette source d'information modifiée.

CHAPITRE 2. ESTIMATEURS DE LA CONSOMMATION POUR UN TRAJET DONNÉ

TABLEAU 2.6 – Comparaison des performances des estimateurs avec information des sous-stations toutes les 10 min et 1 min pour le cas d’étude Bordeaux-Arcachon 2018

écart-type moyen de l’erreur (écart-type de l’écart-type sur 100 estimateurs)	Bordeaux-Arcachon 10 min	Bordeaux-Arcachon 1 min	Arcachon-Bordeaux 10 min	Arcachon-Bordeaux 1 min
Moyenne sur l’ensemble de la base de données	10,33% (0,003%)		11,12% (0,004%)	
Moyenne mensuelle	10,06% (0,010%)		10,69% (0,013%)	
Régression linéaire multiple	9,71% (0,005%)	9,70% (0,005%)	10,54% (0,005%)	10,44% (0,006%)
SVM	9,01% (0,025%)	9,00% (0,026%)	9,46% (0,023%)	9,30% (0,019%)
Réseaux de neurones	9,15% (0,045%)	9,09% (0,050%)	9,59% (0,037%)	9,42% (0,050%)

Si le gain dans les performances est presque nul avec la consommation des sous-stations toutes les minutes, la question de savoir si cette source d’information a pris le relais d’une autre source d’information reste à déterminer. Pour cela, dans le Tableau 2.7, on peut observer l’apport de chaque source d’information avec une consommation de sous-station obtenue toutes les 10 minutes et toutes les minutes.

TABLEAU 2.7 – Comparaison de l’apport de chaque source d’information pour l’estimateur SVM avec information des sous-stations toutes les 10 min et 1 min pour le cas d’étude Bordeaux-Arcachon 2018

Source d’information	Bordeaux-Arcachon 10 min	Bordeaux-Arcachon 1 min	Arcachon-Bordeaux 10 min	Arcachon-Bordeaux 1 min
Jour année	Bruit	8,57 %	10,20 %	14,02 %
Jour semaine	Bruit	Bruit	Bruit	Bruit
Mois	6,82 %	0,41 %	1,20 %	0,42 %
WE/Férié	5,37 %	3,34 %	5,85 %	6,94 %
Heure départ	21,31 %	21,19 %	17,69 %	16,01 %
Durée trajet	14,09 %	13,81 %	Bruit	Bruit
Température	46,36 %	45,82 %	58,14 %	53,04 %
STA 1 - DC	0,96 %	Bruit	Bruit	Bruit
STA 2 - DC	Bruit	Bruit	Bruit	Bruit
STA 3 - DC	Bruit	4,63 %	Bruit	Bruit
STA 4 - DC	5,06 %	2,23 %	6,92 %	9,57 %

On observe que les sources d’information principales (la température ambiante et l’heure de départ) ont toujours le même apport. Cela, en conjonction avec un faible apport, voire négatif des consommations des sous-stations même avec le pas échantillonnage de 1 min, permet de conclure qu’un échantillonnage plus fin sur la consommation des sous-stations n’apporte aucun avantage au moment d’extraire de l’information. À nouveau, la sous-station 4 a été la seule à fournir un apport d’informations considérable, mais comme mentionné plus haut, cela reste un cas particulier difficile à retrouver dans d’autres lignes TER.

2.4.3 Cas d'étude TER n°2 Bordeaux-Dax

Dans ce cas d'étude, 8 sous-stations en DC alimentent la ligne. Les données disponibles, qui s'étalent sur trois années, établissent un total de 3713 trajets Bordeaux-Dax et de 3906 trajets Dax-Bordeaux. Dans le Tableau 2.8, les performances des estimateurs sont montrées, le pourcentage affiché est l'écart-type moyen de l'erreur de 100 estimateurs et le pourcentage affiché entre parenthèses est l'écart-type de l'écart-type sur 100 estimateurs. Après l'essai avec différents nombres de neurones par couche cachée, la configuration optimale trouvée pour ce cas d'étude est une seule couche avec 12 neurones.

TABLEAU 2.8 – Performance des estimateurs de la consommation pour le cas d'étude Bordeaux-Dax

écarts-types moyen de l'erreur (écart-type de l'écart-type sur 100 estimateurs)	Bordeaux-Dax	Dax-Bordeaux
Moyenne sur l'ensemble de la base de données	10,28% (0,004%)	11,80% (0,019%)
Moyenne mensuelle	9,52% (0,013%)	10,74% (0,016%)
Régression linéaire multiple	9,55% (0,071%)	10,69% (0,531%)
SVM	8,66% (0,069%)	10,46% (0,728%)
Réseaux de neurones	9,28% (0,096%)	10,61% (0,704%)

Dans ce cas d'étude, on constate à nouveau que l'estimateur SVM fournit les meilleures performances pour les deux sens de circulation. Précisons néanmoins que dans le cas Dax-Bordeaux, l'écart-type de l'écart-type sur 100 estimateurs est assez élevé pour les estimateurs RLM, SVM et réseaux de neurones. Cela suggère que la base de données pour ce sens de trajet n'est pas très homogène et que des combinaisons particulières de données pourraient produire des estimations de la performance très optimistes ou très pessimistes.

Dans le Tableau 2.9, l'apport de chaque source d'information pour l'estimateur SVM peut être observé. Contrairement au cas précédent, ce n'est pas la température ambiante qui est la source d'information la plus importante, mais le jour de l'année. Cette source d'information, avec l'année de la consommation et l'heure de départ, sont les entrées qui donnent presque tout le gain par rapport à une moyenne. Même si le jour de l'année peut fournir un indice de la température ambiante et donc de la consommation des auxiliaires de la climatisation, l'année et l'heure de départ permettent de modéliser un flux de voyageurs variable dans le temps ainsi que le type de service effectuée (semi-direct ou avec arrêt dans toutes les gares) qui aurait un impact sur la consommation du train. À nouveau, on trouve que l'information de la consommation des sous-stations lors du passage du train n'apporte aucun avantage, ce qui confirme que la signature dans la consommation des TER est diluée par les

CHAPITRE 2. ESTIMATEURS DE LA CONSOMMATION POUR UN TRAJET DONNÉ

consommations des train TGV et FRET qui parcourent la même ligne.

TABLEAU 2.9 – Apport de chaque source d’information pour l’estimateur SVM pour le cas d’étude Bordeaux-Dax

Source d’information	Bordeaux-Dax	Dax-Bordeaux
Jour année	37,99 %	37,75 %
Jour semaine	Bruit	Bruit
Mois	Bruit	Bruit
Année	28,11 %	20,15 %
WE/Férié	Bruit	Bruit
Heure départ	32,78 %	15,50 %
Durée trajet	1,12 %	17,36 %
Température	Bruit	Bruit
STA 1 - DC	Bruit	5,07 %
STA 2 - DC	Bruit	Bruit
STA 3 - DC	Bruit	4,17 %
STA 4 - DC	Bruit	Bruit
STA 5 - DC	Bruit	Bruit
STA 6 - DC	Bruit	Bruit
STA 7 - DC	Bruit	Bruit
STA 8 - DC	Bruit	Bruit

2.4.4 Cas d’étude TER n°3 Dijon-Besançon

Dans ce cas d’étude, 4 sous-stations en DC et 2 sous-stations en AC alimentent la ligne. Les données disponibles, qui s’étalent sur trois ans, établissent un total de 5976 trajets Dijon-Besançon et de 5950 trajets Besançon-Dijon. Dans le Tableau 2.10, les performances des estimateurs sont mises en évidence, le pourcentage affiché est l’écart-types moyen de l’erreur de 100 estimateurs et le pourcentage affiché entre parenthèses est l’écart-type de l’écart-type de ces 100 estimateurs. Après l’essai avec différents nombres de neurones par couche cachée, la configuration optimale trouvée pour ce cas d’étude est une seule couche cachée avec 16 neurones.

TABLEAU 2.10 – Performance des estimateurs de la consommation pour le cas d'étude Dijon-Besançon

écarts-types moyen de l'erreur (écart-type de l'écart-type sur 100 estimateurs)	Dijon-Besançon	Besançon-Dijon
Moyenne sur l'ensemble de la base de données	10,65% (0,003%)	11,39% (0,003%)
Moyenne mensuelle	9,73% (0,009%)	10,36% (0,009%)
Régression linéaire multiple	9,35% (0,075%)	10,05% (0,104%)
SVM	8,52% (0,085%)	9,30% (0,113%)
Réseaux de neurones	8,71% (0,096%)	9,64% (0,104%)

Comme dans les deux cas précédents, c'est l'estimateur SVM qui produit les meilleures performances. Dans ce cas, on trouve un écart-type de l'écart-type de ces 100 estimateurs assez petit pour les deux sens de circulation, ce qui indique que les deux bases de données sont homogènes.

Dans le Tableau 2.11, l'apport de chaque source d'information pour l'estimateur SVM peut être observé. On retrouve à nouveau la température ambiante comme la source d'information la plus importante dans les deux sens de circulation, ce qui souligne la grande consommation des équipements de climatisation. La durée du trajet étant la deuxième source d'information la plus importante, elle pourrait permettre de différencier les types de services TER (avec des arrêts dans toutes les gares ou pas) de cette ligne ainsi que ses consommations. Enfin, l'année de la consommation est une autre source d'information importante, ce qui pourrait indiquer une évolution dans le temps du nombre des voyageurs transportés.

TABEAU 2.11 – Apport de chaque source d’information pour l’estimateur SVM pour le cas d’étude
Dijon-Besançon

Source d’information	Dijon-Besançon	Besançon-Dijon
Jour année	5,15 %	0,93 %
Jour semaine	Bruit	Bruit
Mois	0,85 %	Bruit
Année	18,74 %	15,54 %
WE/Férié	Bruit	Bruit
Heure départ	Bruit	9,74 %
Durée trajet	23,30 %	15,08 %
Température	51,96 %	58,72 %
STA 1 - DC	Bruit	Bruit
STA 2 - DC	Bruit	Bruit
STA 3 - DC	Bruit	Bruit
STA 4 - DC	Bruit	Bruit
STA 5 - AC	Bruit	Bruit
STA 6 - AC	Bruit	Bruit

2.5 Cas d’étude FRET

Pour tous les cas d’étude analysés, l’information de la consommation ainsi que les données d’entrée des estimateurs pour l’année 2020 ont été récupérées. Contrairement aux cas d’étude TER, la faible quantité de données disponibles n’a permis de développer qu’un seul estimateur pour chaque cas d’étude. Le sens de circulation de la ligne a été pris en compte comme une source d’information.

2.5.1 Sources d’information disponibles

Pour les cas d’étude FRET, les sources d’information suivantes sont disponibles :

- Jour de l’année (1 à 365)
- Jour de la semaine (1 à 7)
- Mois (1 à 12)
- WE ou férié (binaire)
- Heure de départ (troncature)
- Durée du trajet (minutes, précision ± 5 min)
- Température à une extrémité du parcours (°C)
- Température à l’autre extrémité du parcours (°C)

- UM (0 si une seule machine, 1 si plusieurs machines)
- TBR, masse remorquée du train, déclaratif (en tonnes)
- TBC, masse totale de train, locomotives incluses, déclaratif (en tonnes)
- Consommation dans les sous-stations lors du passage du train

La durée du trajet a une incertitude de ± 5 min au départ et à l'arrivée en raison de la méthode employée pour déterminer les trajets. Cette méthode utilise en effet les positions GPS données par les compteurs pour déterminer le départ/arrivée du train dans les gares de départ/terminus, mais cette position GPS n'étant récupérable que toutes les 5 min, des incertitudes de ± 5 min sont générées.

Pour les températures ambiantes, elles sont récupérées dans l'aéroport le plus proche des extrémités de la ligne étudiée : l'aéroport de Orly à Paris et l'aéroport de Montélimar pour le cas d'étude Avignon-Valenton, l'aéroport de Perpignan et l'aéroport de Nancy pour le cas d'étude Le Boulou - Zoufftgen et l'aéroport de Lille et l'aéroport de Nancy pour le cas d'étude Dunkerque-Ebange.

Les consommations des sous-stations lors du passage du train ont été reconstruites avec la même méthode que dans les cas d'étude TER. En raison du grand nombre de sous-stations dans chaque cas d'étude, un guide pour décoder le nom de chaque sous-station est disponible en Annexe 2.

Dans tous les cas d'étude FRET, une attention particulière est apportée à l'apport des sources d'information des masses déclarées et des consommations dans les sous-stations lors du passage du train. En effet, ces deux sources d'information sont les plus complexes à récupérer, spécialement la masse déclarée. Une analyse sans aucune de ces deux sources, avec l'information des sous-stations et avec les deux sources a été réalisée.

2.5.2 Cas d'étude FRET n°1 Avignon-Valenton

Dans ce cas d'étude, les trains parcourent une distance totale d'environ 650 km, le tout électrifié avec 1,5kV DC. Dans la Figure 2.2, on peut observer les différents parcours des trains pour le trajet Avignon – Valenton, qui comprend un total de 103 sous-stations. 651 circulations ont été identifiées.

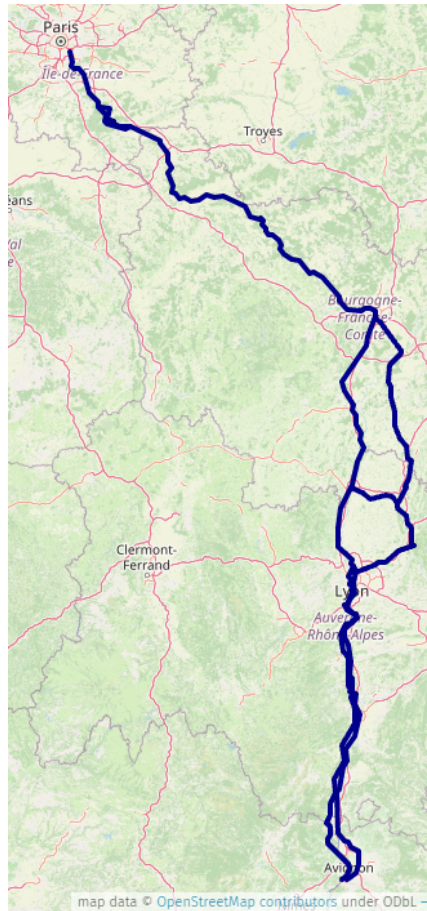


FIGURE 2.2 – Trajet Avignon-Valenton avec les différents parcours

Dans le Tableau 2.12, les performances des différents estimateurs sont indiquées. Dans la première colonne, on donne aux estimateurs toutes les données d'entrée, sauf les consommations dans les sous-stations et la masse du train. Dans la deuxième colonne, on ajoute aux données d'entrée la consommation dans les sous-stations. Dans la troisième colonne, on ajoute l'information de la masse déclarée aux données d'entrée.

TABLEAU 2.12 – Performance des estimateurs de la consommation pour le cas d'étude Avignon-Valenton

écarts-types moyen de l'erreur (écart-type de l'écart-type sur 100 estimateurs)	Avignon-Valenton (651 trajets)	Avignon-Valenton avec sous-stations (651 trajets)	Avignon-Valenton avec sous-stations et masse (651 trajets)
Moyenne sur l'ensemble de la base de données		15,88% (0,025%)	
Moyenne mensuelle		15,04% (0,070%)	
Régression linéaire multiple	13,22% (0,034%)	11,68% (0,062%)	11,48% (0,048%)
SVM	12,07% (0,087%)	13,17% (0,059%)	11,77% (0,072%)

Il apparaît que la meilleure performance est obtenue avec un RLM, les données des sous-stations et la masse du train. Dans ce cas, l'ajout de l'information de la masse des trains n'aide qu'à améliorer légèrement la performance de l'estimateur. Un modèle plus compliqué comme le SVM peut également faire preuve de bonnes performances sans l'information de la consommation des sous-stations. L'utilisation des estimateurs SVM avec les données des sous-stations et la masse n'est pas intéressante mais permet de se rapprocher des performances des RLM.

Dans le Tableau 2.13, on observe que la masse TBR aide à hauteur de 0,24%, ce qui est à peu près la différence de performance entre l'estimateur avec et sans cette donnée. L'heure de départ et la durée de trajet sont les données les plus importantes dans tous les cas. PERRIGNY et GEVREY ont été identifiées comme les sous-stations qui fournissent le plus d'informations pour ce cas d'étude.

TABLEAU 2.13 – Apport des sources d'information les plus importantes pour les estimateurs les plus performants pour le cas d'étude Avignon-Valenton

Avignon-Valenton SVM 12,07 %	Avignon-Valenton avec sous-stations RLM 11,68 %	Avignon-Valenton avec sous-stations et masse RLM 11,48 %
Heure départ 46,99 %	Heure départ 40,71 %	Heure départ 38,86 %
Durée trajet 37,01 %	Durée trajet 22,38 %	Durée trajet 21,36 %
Mois 13,91 %	Sous-station de PERRIGNY 10,24 %	Sous-station de PERRIGNY 10,00 %
Jour année 2,10 %	Sous-station de GEVREY 10,00 %	Sous-station de GEVREY 9,55%
	Sous-station de CEZY 3,33 %	TBR 5,45 %
	Sous-station de AMPUIS 2,14 %	Sous-station de LAROCHÉ 3,41 %

2.5.3 Cas d'étude FRET n°2 Le Boulou - Zoufftgen

Dans ce cas d'étude, les trains parcourent une distance totale d'environ 900 km, dont 300km sont électrifiés avec 25kV AC et le reste avec 1,5kV DC. Dans la Figure 2.3 figurent les différents parcours qu'effectuent les trains pour le trajet Le Boulou – Zoufftgen, lesquels comprennent un total de 105 sous-stations. 334 circulations ont été identifiées.

CHAPITRE 2. ESTIMATEURS DE LA CONSOMMATION POUR UN TRAJET DONNÉ

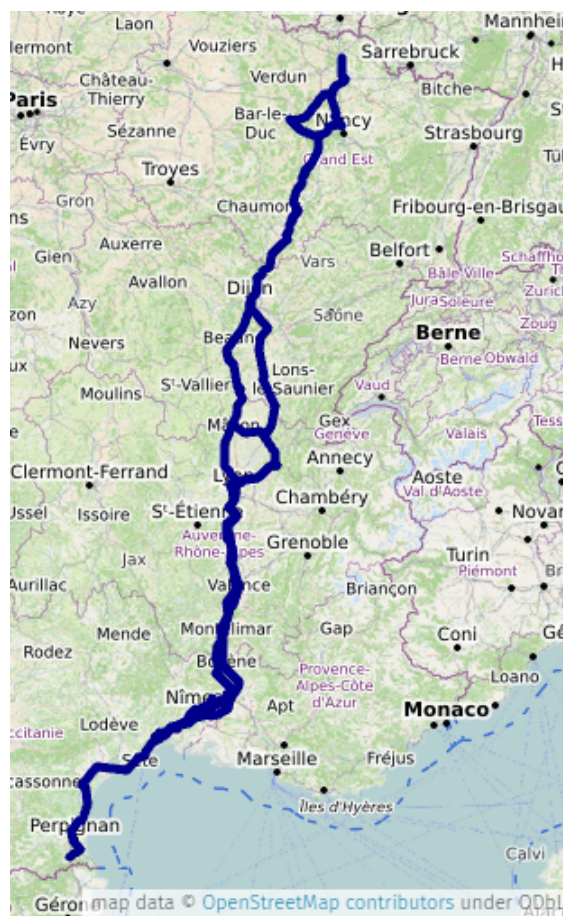


FIGURE 2.3 – Trajet Le Boulou - Zoufftgen avec les différents parcours

Dans le Tableau 2.14 sont indiquées les performances des différents estimateurs. Dans la première colonne, on fournit aux estimateurs toutes les données d'entrée, sauf les consommations dans les sous-stations et la masse du train. Dans la deuxième colonne, on ajoute aux données d'entrée la consommation dans les sous-stations. Dans la troisième colonne, on ajoute l'information de la masse déclarée aux données d'entrée.

TABLEAU 2.14 – Performance des estimateurs de la consommation pour le cas d'étude Dunkerque-Ebange

écarts-types moyen de l'erreur (écart-type de l'écart-type sur 100 estimateurs)	Le Boulou-Zoufftgen (334 trajets)	Le Boulou-Zoufftgen avec sous-stations (334 trajets)	Le Boulou-Zoufftgen avec sous-stations et masse (334 trajets)
Moyenne sur l'ensemble de la base de données		35,33% (0,116%)	
Moyenne mensuelle		27,70% (0,465%)	
Régression linéaire multiple	29,17% (0,185%)	26,29% (0,194%)	26,39% (0,187%)
SVM	26,72% (0,244%)	28,99% (0,286%)	28,44% (0,333%)

Dans ce cas d'étude en particulier, l'amélioration entre la moyenne annuelle et mensuelle est sensible, ce qui témoigne d'un effet « saison » très important. On peut supposer qu'il s'agit des effets du COVID-19 et les confinements qui s'en sont suivi. Il ressort que la meilleure performance est obtenue avec un RLM et les données de sous-stations. L'ajout de l'information de la masse des trains n'aide pas à améliorer la performance de l'estimateur. On remarque en outre qu'un modèle plus compliqué comme le SVM peut aussi produire de bonnes performances sans l'information de la consommation des sous-stations. L'utilisation des estimateurs SVM avec les données des sous-stations et masse n'est pas intéressante. Cela pourrait s'expliquer par le faible nombre de trajets disponibles. Dans le Tableau 2.15, on observe que comme les masses TBC et TBR n'aident qu'à hauteur de 0,10% et 0,06% respectivement. On en conclut que pour ce cas d'étude, la donnée de la masse déclarée n'est pas une source d'information significative. Le jour de l'année est en revanche la donnée la plus importante lorsque les données des sous-stations ne sont pas disponibles. Le nombre de machines du train (UM) est pour sa part la donnée la plus importante quand on dispose des données des sous-stations, et la deuxième en importance quand celles-ci ne sont pas disponibles. NIMES a été identifiée comme la sous-station qui fournit le plus d'informations pour ce cas d'étude. Finalement, on constate que toutes les sous-stations identifiées comme source d'information, sont des sous-stations DC. Aucune sous-station AC n'a été identifiée comme source d'information.

TABLEAU 2.15 – Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Le Boulou-Zoufftgen

Le Boulou-Zoufftgen SVM 26,72%	Le Boulou-Zoufftgen avec sous-stations RLM 26,29%	Le Boulou-Zoufftgen avec sous-stations et masse RLM 26,39%
UM 93,61 %	UM 59,16 %	UM 60,16 %
Direction 5,11 %	STA 69 9,66 %	STA 69 9,46 %
Heure 1,28 %	STA 88 3,88 %	STA 88 3,62 %
Jour semaine 0,81 %	STA 56 2,59 %	STA 94 2,21 %
	STA 89 1,99 %	STA 89 1,91 %
	STA 81 1,89%	STA 56 1,81 %

2.5.4 Cas d'étude FRET n°3 Dunkerque - Ebange

Dans ce cas d'étude, les trains parcourent une distance totale d'environ 400 km, le tout électrifiée avec 25kV AC. Dans la Figure 2.4 sont indiqués les différents parcours qu'effectuent les trains pour le trajet Dunkerque - Ebange, lesquels comprennent un total de 15 sous-stations. 733 circulations ont été identifiées.



FIGURE 2.4 – Trajet Dunkerque - Ebange avec les différents parcours

Dans le Tableau 2.16 figurent les performances des différents estimateurs. Dans la première colonne, on fournit aux estimateurs toutes les données d'entrée, sauf les consommations dans les sous-stations et la masse du train. Dans la deuxième colonne, on ajoute aux données d'entrée la consommation dans les sous-stations. Dans la troisième colonne, on ajoute l'information de la masse déclarée aux données d'entrée.

Il apparaît que la meilleure performance est obtenue avec un SVM et les données des sous-stations. Dans ce cas, l'ajout de l'information de la masse des trains n'aide pas à améliorer la performance de

l'estimateur. Les estimateurs SVM sont toujours plus performants que les estimateurs RLM, ce qui s'explique par le nombre de trajets disponibles plus important que dans les cas d'étude précédents.

TABLEAU 2.16 – Performance des estimateurs de la consommation pour le cas d'étude Dunkerque-Ebange

écarts-types moyen de l'erreur (écart-type de l'écart-type sur 100 estimateurs)	Dunkerque-Ebange (733 trajets)	Dunkerque-Ebange avec sous-stations (733 trajets)	Dunkerque-Ebange avec sous-stations et masse (733 trajets)
Moyenne sur l'ensemble de la base de données	17,90% (0,043%)		
Moyenne mensuelle	15,32% (0,105%)		
Régression linéaire multiple	15,27% (0,047%)	14,69% (0,060%)	14,71% (0,057%)
SVM	14,15% (0,126%)	11,72% (0,146%)	11,77% (0,130%)

Dans le Tableau 2.17, on observe que la masse TBC aide à hauteur de 0,03%. On peut donc en conclure que pour ce cas d'étude, la donnée de la masse déclarée n'est pas une source d'information importante. Dans ce cas, l'ajout de l'information des consommations des sous-stations aide grandement à améliorer la performance. Contrairement à ce qu'on aurait pu penser, l'information peut être extraite des consommations des sous-stations AC.

TABLEAU 2.17 – Apport des sources d'information le plus importantes pour les estimateurs les plus performants pour le cas d'étude Dunkerque-Ebange

Dunkerque-Ebange SVM 14,15%	Dunkerque-Ebange avec sous-stations SVM 11,72%	Dunkerque-Ebange avec sous-stations et masse SVM 11,77%
Durée trajet 46,40 %	Sous-station de VERQUIGNEUL 27,67 %	Sous-station de THIONVILLE 24,31 %
Heure départ 32,53 %	Sous-station de THIONVILLE 13,92 %	Température à Lille 14,36 %
Température à Lille 6,67 %	Sous-station de MAING 12,78 %	Sous-station de VERQUIGNEUL 14,03 %
Jour semaine 4,53 %	Sous-station de JOULDREVILLE 12,46 %	Sous-station de MAING 12,89 %
Mois 4,00 %	Mois 11,97 %	Sous-station de BUSIGNY 12,07 %
Température à Nancy 3,47 %	Heure départ 7,12 %	Jour année 10,60 %

2.6 Conclusion

Dans les cas d'étude TER, les estimateurs du type SVM ont les meilleures performances avec en contrepartie un incrément dans le temps de calcul en comparaison avec les moyennes. Les réseaux de neurones, n'arrivent jamais à dépasser les performances des estimateurs SVM, de plus le temps de calcul nécessaire pour créer ces estimateurs est très grande ce qui rend les réseaux de neurones peu attractifs. Dans tous les cas, on a pu constater que l'amélioration obtenue par rapport à une moyenne est très limitée. En effet, les sources d'information disponibles ne témoignent pas d'une grande corrélation avec la consommation des trains. Néanmoins, comme il peut être observé dans le Tableau 2.18, quelques sources se sont avérées importantes : la température ambiante pour les cas d'étude Bordeaux-Arcachon et Dijon-Besançon est une donnée responsable de plus de la moitié de l'amélioration de la performance des estimateurs. Le jour de l'année, l'année de la consommation et l'heure de départ du train sont également identifiés comme des sources d'information importantes pour modéliser l'évolution du nombre de passagers et donc la consommation du train. Dans tout les cas, le jour de la semaine a été identifié comme une source de bruit, ce qui implique qu'une grande variation du nombre de voyageurs n'est pas présente sur ces lignes étudiées ou que cette variation peut être modélisée à partir les autres sources d'information. Les consommations des trains TER sont toutefois trop faibles pour que leur signature soit récupérée dans les consommations relevées dans les sous-stations. En effet, cette signature est effacée par des consommations plus élevées comme celles des trains TGV ou FRET.

TABLEAU 2.18 – Apport moyen de chaque source d'information pour l'estimateur SVM pour les cas d'étude TER

Source d'information	Apport moyen d'information dans les cas d'étude TER
Jour année	16,33 %
Jour semaine	Bruit
Mois	1,54 %
Année	19,20 %
WE/Férié	1,15 %
Heure départ	14,60 %
Durée trajet	10,56 %
Température	32,58 %
Conso sous-stations tous apports confondues	7,48 %

Concernant les cas d'étude FRET, la première conclusion est que la donnée de la masse n'améliore généralement pas la performance des estimateurs quand les données des sous-stations sont disponibles. Cela peut être dû au caractère déclaratif de ce type de données ainsi que la faible variation de cette

donnée pour certains cas d'étude (ex. Le Boulou - Zoufftgen) ou au fait que le gain apporté par les masses peut être extrait des consommations des sous-stations. Comme il peut être observé dans le Tableau 2.19, il n'y a pas une source d'information qui s'impose clairement sur les autres, néanmoins l'heure départ et l'UM sont des sources d'information relativement importantes. C'est l'information provenant des sous-stations qui est la plus importante, toutefois cette information est répartie entre un grand nombre de sous-stations et l'apport d'information de chaque sous-station est donc faible. La deuxième conclusion est que les performances des estimateurs RLM avec les données des sous-stations peuvent être atteintes avec des estimateurs SVM sans les données des sous-stations. Ce n'est que si un grand nombre de données est disponible que l'intérêt des estimateurs SVM avec des données des sous-stations est évident. En moyenne, un gain de 6,54 % dans les performances des estimateurs plus adaptés est obtenu en comparaison avec une moyenne annuelle.

TABLEAU 2.19 – Apport moyen de chaque source d'information pour le meilleur estimateur pour les cas d'étude FRET

Source d'information	Apport moyen d'information dans les cas d'étude FRET
Jour année	0,38 %
Jour semaine	Bruit
Mois	3,00 %
WE/Férié	0,02 %
Heure départ	15,73 %
Durée trajet	7,09 %
Direction	Bruit
Température tous apports confondues	2,42 %
UM	21,92 %
Masse déclarée	2,18 %
Conso sous-stations DC tous apports confondues	59,55 %
Conso sous-stations AC tous apports confondues	88,07 %

Dans le futur, il pourrait être intéressant de récupérer les consommations avec un pas d'échantillonnage de 1 min dans les sous-stations. Cet échantillonnage plus fin pourrait permettre d'extraire plus d'informations, spécialement dans les sous-stations DC (où le changement de sous-station est plus fréquent qu'en AC).

En vue de la disparité des résultats obtenus, l'établissement d'une méthodologie globale pour tous les cas d'étude semble difficile. Néanmoins, quelques recommandations peuvent être formulées. Pour le cas d'étude TER, les estimateurs du type SVM devrait être l'option privilégiée, car il est l'estimateur qui arrive à extraire le plus des sources d'information tout en ayant un temps de calcul modéré.

Concernant la pertinence des sources d'information, la température ambiante ainsi que l'année et le jour de l'année ont été identifiées comme les sources d'information le plus importantes. Néanmoins, une étude cas par cas devrait être menée pour vérifier cette dernière information, parce que dans des cas concrets cette sources peuvent même devenir de sources de bruit (e.g. la température ambiante pour le cas d'étude Bordeaux-Dax). Dans le cas d'étude FRET, le choix d'un type d'estimateur est plus compliquée. Les estimateurs RLM ont été identifiés comme les estimateurs le plus adaptés dans deux cas d'études tandis que les estimateurs SVM l'ont été pour le troisième cas d'étude, donc une étude comparative cas par cas doit être fait pour déterminer l'estimateur le plus adapté. En regardant les sources d'information, la disparité est encore plus grande, dans certains cas d'étude une seule source d'information peut apporter presque toute l'amélioration de l'estimateur (comme l'information sur les UM dans le cas d'étude Le Boulou-Zoufftgen ou l'heure de départ et la durée du trajet pour le cas d'étude Avignon-Valenton). Donc, à nouveau une analyse cas par cas doit être faite, on peut remarquer néanmoins que la récupération des consommations des sous-stations AC semble plus pertinente que celles des sous-stations DC pour apporter plus d'information plus concentrée car provenant de moins de sous-stations.

Chapitre 3

Impact des smart-grids ferroviaires dans le comptage d'énergie

Contenu

2.1	Introduction	56
2.2	Estimateurs « data-based »	57
2.2.1	Moyenne sur l'ensemble de la base de données	58
2.2.2	Moyenne mensuelle	58
2.2.3	Régression linéaire multiple (RLM)	58
2.2.4	Support Vector Machine (SVM)	58
2.2.5	Réseaux de neurones	59
2.3	Méthode pour trier les sources d'information disponibles	59
2.4	Cas d'étude TER	60
2.4.1	Sources d'information disponibles	60
2.4.2	Cas d'étude TER n°1 Bordeaux-Arcachon	61
2.4.3	Cas d'étude TER n°2 Bordeaux-Dax	65
2.4.4	Cas d'étude TER n°3 Dijon-Besançon	66
2.5	Cas d'étude FRET	68
2.5.1	Sources d'information disponibles	68
2.5.2	Cas d'étude FRET n°1 Avignon-Valenton	69
2.5.3	Cas d'étude FRET n°2 Le Boulou - Zoufftgen	71
2.5.4	Cas d'étude FRET n°3 Dunkerque - Ebange	74
2.6	Conclusion	76

3.1 Introduction

Les chemins de fer représentent l'alternative la plus respectueuse pour l'environnement s'agissant du transport des personnes et des marchandises. Principalement électrique, leur nature garantit leur efficacité et des émissions de CO₂ très faibles, voire nulles. D'après la SNCF, un voyage en TGV est 40 fois moins polluant qu'en voiture particulière et 70 fois moins polluant qu'en avion [37]. En raison de son efficacité et de sa moindre consommation par rapport au trafic routier, de nouvelles combinaisons de train plus camion (appelées autoroutes ferroviaires) commencent à apparaître en Europe [38]. Le fait que le transport ferroviaire soit électrique implique qu'il dépend du mix électrique de chaque pays en termes d'émissions de CO₂. Différentes EF comme le opérateur néerlandais NS (Nederlandse Spoorwegen) [39] ou le train à grande vitesse de Californie [40] ont pour objectif de s'alimenter seulement avec une énergie d'origine 100 % renouvelable. Au niveau national, des projets comme CONIFER (Conception et Outils iNnovants pour un réseau électrique Intelligent appliqué au FERroviaire) s'intéressent aussi à l'intégration des sources d'énergie renouvelable dans le domaine ferroviaire [41]. C'est à ce niveau qu'interviennent les smart-grids : dans [42] et [43], ce type de réseau intègre directement la production dans le système ferroviaire. La notion de « sous-station hybride » est mentionnée dans ces ouvrages où, en plus d'intégrer les sources d'énergie renouvelable, du stockage d'énergie est rajouté pour favoriser la récupération d'énergie de freinage sous caténaire DC. Une autre option mentionnée dans la littérature pour faciliter cette récupération d'énergie est d'installer un convertisseur dans la sous-station en la rendant réversible afin de pouvoir réinjecter l'énergie dans le réseau [44] [45].

L'introduction de ce type de smart-grids dans le système ferroviaire pourrait impacter la consommation totale des trains, sa distribution et les estimateurs de l'énergie consommée développés dans cette thèse.

3.2 Impact des smart-grids dans la consommation des trains

Deux cas ont été analysés. Dans le premier, la récupération sous DC serait comparable à la récupération sous AC existante aujourd'hui et dans le deuxième, les trains participeraient à l'effacement de la consommation sur demande.

3.2.1 Analogie entre les électrifications en DC et AC vis-à-vis de l'énergie récupérée

L'énergie récupérée lors des phases de freinage est naturellement réinjectée dans la caténaire puis dans le réseau de distribution d'électricité sous caténaire AC. Néanmoins, sous caténaire DC, en raison de l'irréversibilité des sous-stations, l'énergie récupérée ne peut être consommée que par les trains des alentours, en limitant par ailleurs la quantité d'énergie récupérable à l'énergie absorbée par les trains présents. L'introduction des smart-grids, sous la forme de sous-stations en DC réversibles ou avec du stockage, peut faire qu'un système d'alimentation en DC adopte le même comportement qu'un système d'alimentation AC concernant la récupération de l'énergie de freinage. Cela pourrait entraîner la diminution de la consommation totale du train, tout en réduisant l'impact du style de conduite sur la consommation en limitant la dispersion des consommations. Ces hypothèses vont être testées à travers un cas d'étude TER et un autre FRET. Dans les deux cas, la ligne dispose de sous-stations AC et DC. Un train avec les mêmes caractéristiques et conditions (même masse, conditions climatologiques, conducteur...) passera par les deux types d'alimentations et sa consommation fera apparaître les particularités de chaque type d'alimentation.

3.2.1.1 Cas d'étude TER

Les trains de type TER, dont les arrêts sont fréquents, ont un haut potentiel de récupération lors de phases de freinage. Pour étudier cette caractéristique, nous avons sélectionné la ligne TER Dijon-Besançon. Cette ligne est électrifiée en courant continu entre Dijon et Dole (45,88 km) et en courant alternatif entre Dole et Besançon (45,24 km). La Figure 3.1 présente l'histogramme de l'énergie récupérée lors des freinages sous caténaire DC. La moyenne pour le sens Dijon-Besançon est de 1,97 % avec un écart-type de 2,28 %. Pour le sens Besançon-Dijon, la moyenne est de 2,91 % avec un écart-type de 2,59 %.

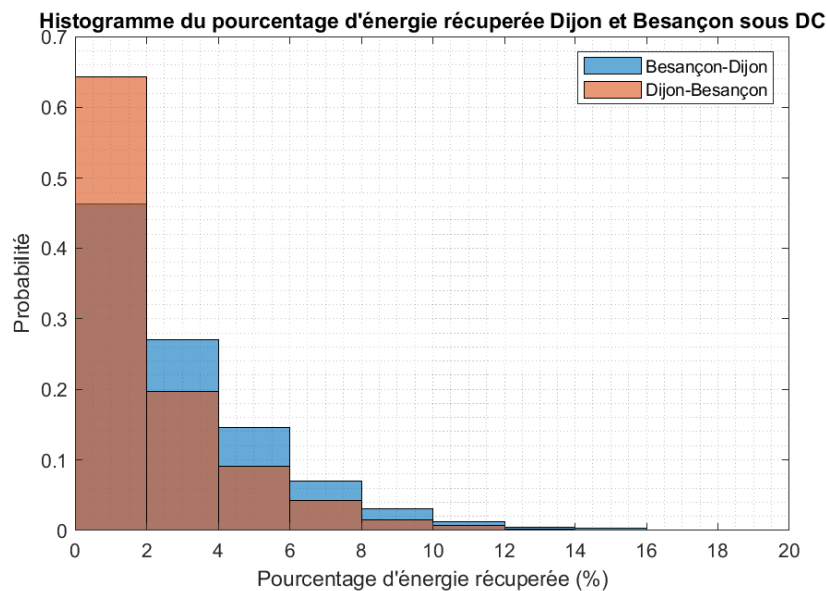


FIGURE 3.1 – Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation DC

Dans la Figure 3.2 est présenté l'histogramme de l'énergie récupérée lors des freinages sous caténaire AC. La moyenne dans le sens Dijon-Besançon est de 21,71 % avec un écart-type de 6,31 %. Pour le sens Besançon-Dijon, la moyenne est de 18,48 % avec un écart-type de 7,77 %. Il apparaît que la quantité d'énergie récupérée sous caténaire AC est bien plus grande qu'en DC, ce qui confirme l'hypothèse selon laquelle avec des sous-stations DC «hybrides» ou réversibles, la consommation moyenne des trains devrait être considérablement réduite. Néanmoins, les figures 3.1 et 3.2 ne permettent pas de déterminer si la dispersion des consommations totales est différente suivant le type d'alimentation.

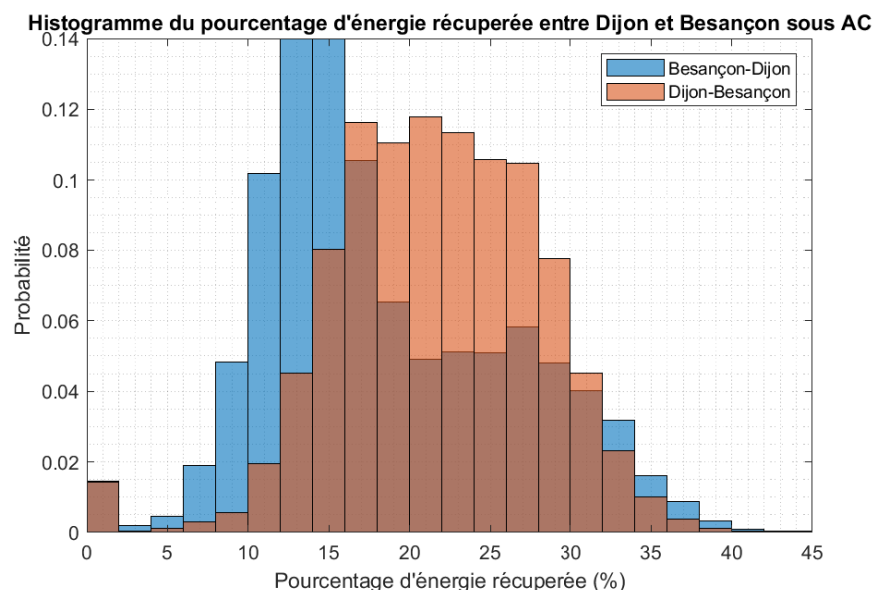


FIGURE 3.2 – Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation AC

Dans les Figures 3.3 et 3.4 apparaissent les consommations totales (énergie consommée moins l'énergie récupérée) sous les deux types d'alimentation pour les trajets Dijon-Besançon et Besançon-Dijon respectivement. L'écart-type est la valeur prise en compte pour mesurer la dispersion des consommations et donc l'effet de la conduite sur la consommation. Pour le sens Dijon-Besançon, l'écart-type est de 15,76 % pour les consommations AC et de 15,07 % pour les consommations DC. Pour le sens Besançon-Dijon, l'écart-type est de 13,75 % pour les consommations AC et de 13,85 % pour les consommations DC. On en conclut que contrairement à l'hypothèse émise, la récupération d'énergie de freinage n'affecte pas la dispersion des consommations. Le fait d'introduire des sous-stations DC «hybrides» ou réversibles dans les alimentations DC fera descendre la valeur moyenne d'énergie consommée, mais n'impactera pas sa dispersion dans le cas de trains de type TER.

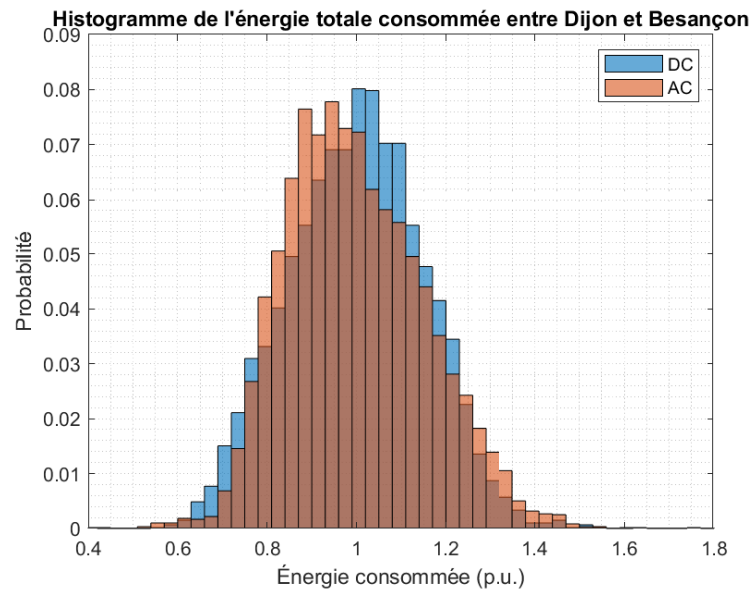


FIGURE 3.3 – Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets
Dijon-Besançon

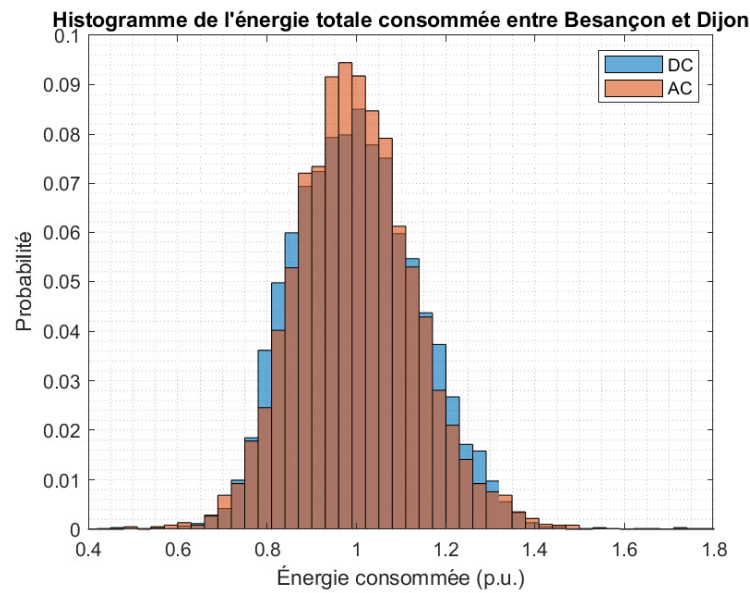


FIGURE 3.4 – Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets
Besançon-Dijon

3.2.1.2 Cas d'étude FRET

Les trains FRET marquent moins d'arrêts que les TER, néanmoins, lors des descentes, la quantité d'énergie qui peut être récupérée n'est pas négligeable. La ligne Le Boulou-Zoufftgen a été sélectionnée parce qu'elle est alimentée avec les deux types d'électrification comme dans le cas d'étude TER précédent. À titre de rappel, 300km sont électrifiés en 25kV AC et 600km en 1,5kV DC. Dans la Figure 3.5 sont indiqués les pourcentages d'énergie récupérée sous caténaire DC. Pour le sens Le Boulou-Zoufftgen, la moyenne est de 0,74 % avec un écart-type de 0,71 % et pour le sens Zoufftgen-Le Boulou, de 1,02 % avec un écart-type de 0,85 %. On en déduit à nouveau que l'énergie récupérée sous DC est très faible avec l'infrastructure actuelle.

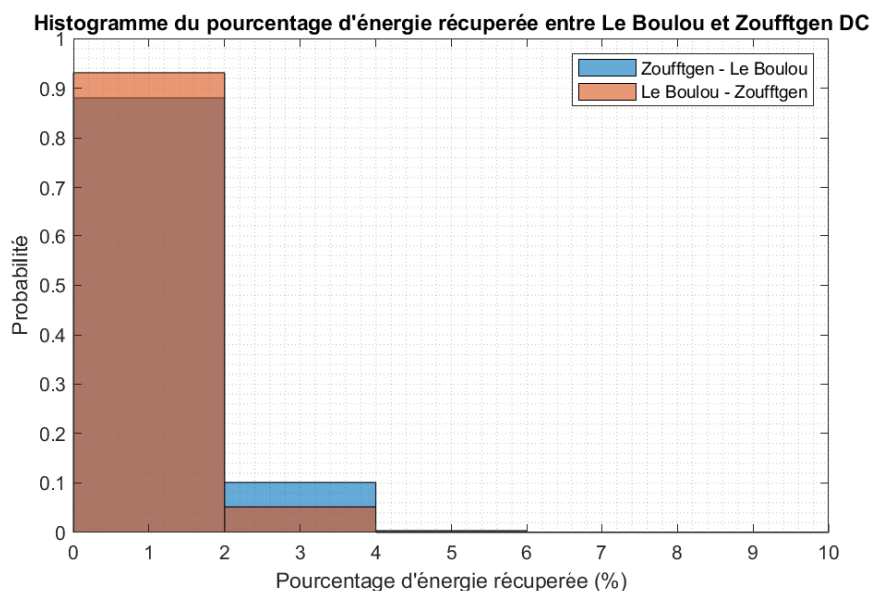


FIGURE 3.5 – Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation DC

Dans la Figure 3.6 apparaissent les pourcentages d'énergie récupérée sous caténaire AC. La moyenne d'énergie récupérée dans les zones AC est de 12,74 % avec un écart-type de 4,17 % pour le sens Le Boulou-Zoufftgen et de 12,45 % avec un écart-type de 3,77 % dans le sens inverse. Le pourcentage d'énergie récupérée est bien inférieur dans le cas des trains FRET que dans le cas des trains TER, mais au regard de la grande consommation des trains FRET, ce petit pourcentage implique une quantité d'énergie non négligeable. Dans ce cas d'étude, comme dans le précédent, c'est principalement dans les zones électrifiées sous AC que l'énergie est récupérée.

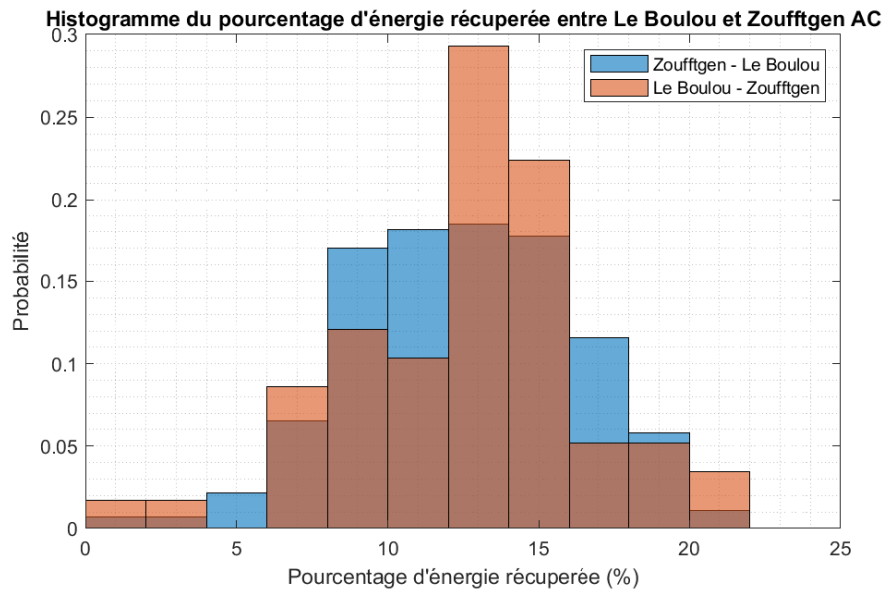


FIGURE 3.6 – Histogramme du pourcentage d'énergie récupérée par rapport à l'énergie consommée en alimentation AC

Pour évaluer l'éventuel impact des sous-stations DC «hybrides» ou réversibles sur la dispersion des consommations, on fait à nouveau référence aux consommations totales. Il ressort des Figures 3.7 et 3.8 que la dispersion est sensiblement identique sous les deux types d'alimentation. Dans le sens Zoufftgen–Le Boulou, l'écart-type sous caténaire AC est de 45,77 % et de 35,69 % sous caténaire DC. Dans le sens inverse, l'écart-type sous caténaire AC est de 25,63 % et de 27,04 % sous caténaire DC.

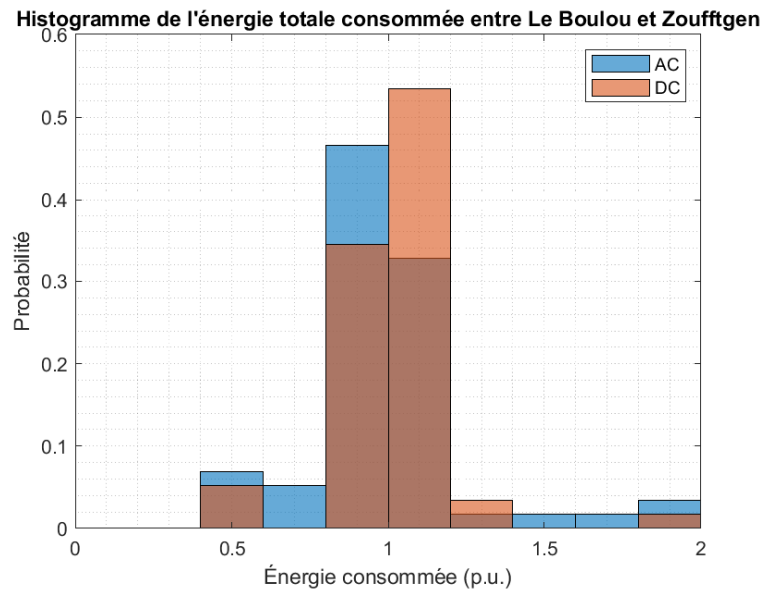


FIGURE 3.7 – Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets Le Boulou – Zoufftgen

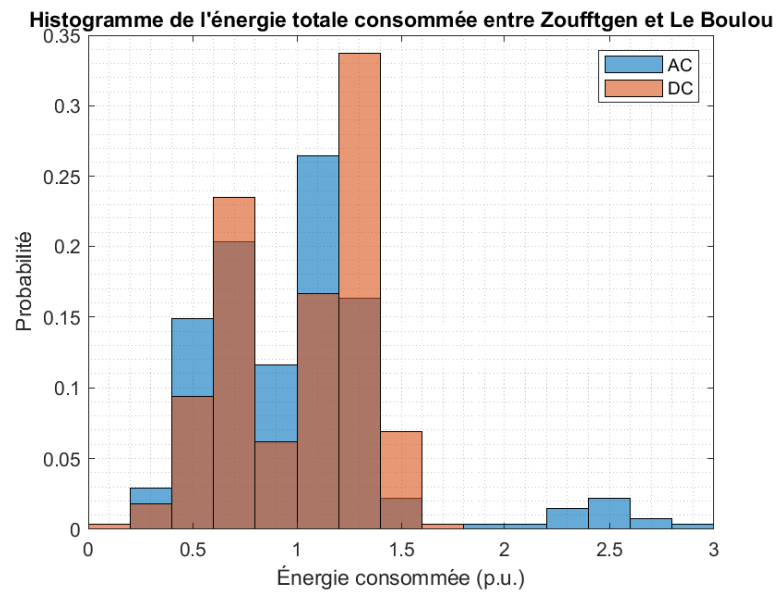


FIGURE 3.8 – Histogramme de la consommation totale sous les deux types d'alimentation pour les trajets Zoufftgen-Le Boulou

3.2.2 Effacement de la consommation ferroviaire sur demande

L'effacement de la consommation ferroviaire est un autre type de fonctionnalité développée dans un smart-grid ferroviaire qui pourrait être mis en place et qui impacterait la consommation des trains [46]. Cet effacement a pour but d'éviter de solliciter le réseau public d'électricité dans les moments où la consommation sur le réseau est la plus forte. Dans le domaine ferroviaire, cet effacement peut être assuré de deux manières : par une réduction de la vitesse de consigne des trains et par la variation de la consommation des organes de confort (chauffage, ventilation et climatisation) [47].

La réduction de la vitesse consigne est une méthode adaptée pour des trains où une grande quantité d'énergie cinétique est stockée, soit principalement des trains TGV. Dans la Figure 3.9, en rouge, on observe un parcours normal sans période d'effacement. En bleu, le trajet affiche une période d'effacement, mais le retard est compensé en consommant plus avant et après cette période. Finalement, en vert, le trajet est marqué par une période d'effacement et un retard, ce qui implique une consommation inférieure à un trajet standard.

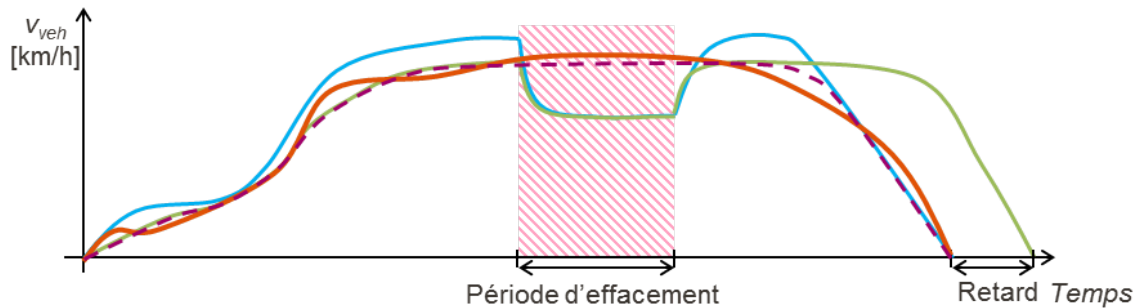


FIGURE 3.9 – Différents profils de vitesse pour un trajet donné sans effacement en rouge, avec effacement et sans retard en bleu et avec effacement et retard en vert [46]

La variation de la consommation des organes de confort visant quant à elle les trains avec des voyageurs debout, qui sont ainsi moins sensibles aux variations de température, soit les trains TER et RER. Dans ce cas, la puissance des équipements est limitée pendant les périodes d'effacement, ce qui a peu d'incidence sur le confort des voyageurs, mais un impact notable sur la consommation du train.

Si finalement, l'effacement de la consommation est mis en place, au-delà de l'effet sur la consommation moyenne des trains, un impact sur la dispersion des consommations est probable en augmentant cette dispersion. Néanmoins, si les trains qui participent aux périodes d'effacement sont identifiables, l'impact négatif sur les estimateurs pourrait être limité en utilisant cette participation comme une source d'information.

3.3 Impact des smart-grids dans les estimateurs de la consommation

L'introduction des sous-stations en DC réversibles ou avec de la capacité de stockage va « perturber » la quantité et la qualité de l'information qu'on peut obtenir des lectures des compteurs installées dans les sous-stations à date d'aujourd'hui. Avant l'intégration des smart-grids ferroviaires, les sous-stations (et donc ses compteurs) étaient en effet le seul point d'injection d'énergie dans le réseau ferroviaire. La signature de la consommation d'un train n'était donc présente que dans ces compteurs. À l'avenir, avec l'intégration des smart-grids ferroviaires, d'autres points d'injection d'énergie apparaîtront et généreront des flux d'énergie non perçus par les compteurs des sous-stations. Dès lors, l'information alimentée par les données qui pouvait jusqu'à présent être extraite des seuls relevés de compteurs des sous-stations sera encore plus difficile à extraire. Cette perte d'information n'affecte toutefois pas les cas d'étude TER, parce que, comme démontré précédemment (section 2.4), les compteurs des sous-stations n'ont dans aucun cas été identifiés comme une source d'information. Pour les cas d'étude FRET, si l'on envisage le pire cas où l'information des compteurs serait tellement « bruitée » qu'elle deviendrait inexploitable, la perte dans la performance des estimateurs serait comprise entre 0,6 % et 2,4 % selon le cas d'étude, comme on peut l'observer dans le Tableau 3.1. La perte de cette source d'information souligne l'intérêt des modèles plus complexes quand peu d'entrées sont disponibles.

TABLEAU 3.1 – Comparaison des performances des estimateurs dans les cas d'étude FRET avec et sans les consommations dans les sous-stations comme source d'information

Performance des estimateurs (écart-type de l'erreur)	Le Boulou - Zoufftgen	Avignon - Valenton	Dunkerque-Ebange
Meilleure performance avec l'information des sous-stations	RLM 26,29 %	RLM 11,48 %	SVM 11,72 %
Meilleure performance sans l'information des sous-stations	SVM 27,40 %	SVM 12,06 %	SVM 14,14 %

On peut dès lors conclure que l'impact des smart-grids dans les estimateurs de la consommation des trains type TER est nul. Pour les trains type de FRET en revanche, un impact est observé dans la performance des estimateurs, même s'il peut être limité en choisissant des estimateurs plus complexes ou si l'on passe d'estimateurs simples comme les RLM à des estimateurs plus complexes comme les SVM.

3.4 Conclusion

Dans ce chapitre, l'impact de l'évolution du réseau ferroviaire vers un smart-grid ferroviaire sur la consommation ferroviaire et le comptage de l'énergie est discuté. L'analyse de données réalisée à travers des cas d'étude où les deux types d'électrification sont présentes a révélé que l'introduction des sous-stations en DC réversibles ou avec de la capacité de stockage répercutera sur la valeur moyenne de l'énergie consommée par un train, sans toutefois impacter la dispersion de ces consommations. L'introduction d'un effacement de la consommation dans le cadre ferroviaire pourrait quant à elle bien augmenter la dispersion des consommations. L'identification des trains participant à cet effacement et l'utilisation de cette information comme une entrée dans les estimateurs de la consommation limiteraient cependant l'impact sur la performance de ces derniers.

Conclusion et perspectives

Cette thèse apporte de nouvelles méthodes pour faire une validation des données de comptage enregistrées dans les compteurs embarqués dans les trains, auparavant non observés dans la littérature. Elle montre aussi le développement des estimateurs de la consommation pour un trajet et un train donnés, en complétant les estimateurs proposés dans la littérature qui estiment la consommation à l'échelle d'une flotte des trains, disponibles dans la littérature. Le classement par importance des différentes sources d'information enrichisse les propositions de la littérature.

Dans le premier chapitre, une méthode générant des bornes de validation pour évaluer la consommation des trains a été présentée afin d'augmenter la fiabilité des données de comptage. Basées sur des modèles physiques, ces bornes permettent d'écarter les mesures aberrantes du point de vue physique des compteurs embarqués. Si la méthode utilisée pour générer la borne supérieure de validation correspondant au « maximum ferroviaire » peut être utilisée pour des cas d'étude TER et FRET, celle visant à générer la borne inférieure de validation n'est pas adaptée pour des cas d'étude FRET. Néanmoins, si à l'avenir, la position et la vitesse d'un train FRET à des moments donnés et à des intervalles d'une dizaine de minutes pourraient être déterminées, la méthode pourrait également s'appliquer aux cas des trains FRET. De même, si la position et la vitesse des trains TER étaient connues à des intervalles réguliers, le processus de calcul des bornes inférieures serait plus rapide avec des bornes plus serrées. D'autres modélisations mathématiques ou méthodes de résolution peuvent être envisagés si finalement la position et la vitesse d'un train à des moments donnés et à des intervalles d'une dizaine de minutes ne peuvent pas être récupérées.

Dans le deuxième chapitre de la thèse, divers estimateurs de la consommation ont été analysés. Tous sont issus d'une base de données constituée par des mesures de comptage des compteurs embarqués validées par la méthode du premier chapitre. La différence trouvée dans la performance entre les estimateurs les plus simples (comme la moyenne sur toute la base de données ou la moyenne mensuelle) et les plus complexes (RLM, SVM et réseau de neurones) s'explique par l'intégration de sources d'information diverses. Parmi elles, on trouve principalement des données calendaires, des horaires, des données météorologiques et des consommations des sous-stations de traction. Dans le cas d'étude FRET, la masse déclarée du train est une source d'information additionnelle, qui n'est pas pertinente

dans le cas d'étude TER. Les résultats ont révélé que la source d'information la plus importante dans les cas d'étude TER est la température ambiante suivie par certaines données calendaires ou des horaires. Les consommations des sous-stations n'ont presque jamais été identifiées comme une source d'information utile. Néanmoins, aucun ordre d'importance des sources d'informations généralisable ne semble se profiler et une analyse au cas par cas est nécessaire. Dans les cas d'étude FRET, la masse déclarée, d'une façon inattendue, n'a pas été identifiée comme une source d'information importante, ce qui pourrait refléter un écart important entre la masse déclarée et la masse réelle du train en circulation. S'agissant des consommations des sous-stations, c'est quand la consommation provient des sous-stations AC qu'une information utile peut être extraite. Ces résultats ont été obtenus avec des mesures prises toutes les 10 min. Il faudrait désormais tester si avec des consommations relevées toutes les minutes, les sous-stations de DC pourraient également devenir des sources d'information. En raison d'une moindre distance moyenne entre deux sous-stations en DC en comparaison aux sous-stations AC, le pas de mesure de 10 min pour les consommations est trop grand, ce qu'il dilue la signature de la consommation du train avec les autres consommations. Ces estimateurs développés ont pour objectif de reconstruire une consommation qui a déjà eu lieu, néanmoins, avec quelques adaptations, les estimateurs pourraient établir des prévisions de la consommation. Dans le cas d'étude TER par exemple, il suffirait d'anticiper la température ambiante pour prévoir la consommation d'une circulation. Pour les cas d'étude FRET, la consommation des sous-stations lors du passage du train étant une source d'information utile, la faisabilité d'un estimateur des consommations futures est moins pertinente. Quoi qu'il en soit, comme on l'a vu dans le Tableau 3.1, avec des modèles plus complexes, on peut obtenir des estimations proches sans l'information des sous-stations. D'autres types d'estimateurs non analysés dans cette thèse (e.g. les réseaux de neurones d'apprentissage profond) pourraient être envisagées pour estimer la consommation.

Dans le dernier chapitre, l'impact des smart-grids sur la consommation des trains et sur les estimateurs a été analysé. Les smart-grids ont généralement un impact important sur la consommation moyenne des trains, mais n'ont aucune répercussion sur sa dispersion même si parfois, ils peuvent l'augmenter. Vis-à-vis des estimateurs, l'impact le plus important serait la perte des consommations des sous-stations comme source d'information. Dans les cas TER, elles n'étaient pas des sources d'information donc l'impact sur les estimateurs est nul. Dans les cas FRET, cet impact est un peu plus significatif même s'il peut être amorti en choisissant un estimateur plus complexe.

À partir des outils et méthodes développés dans cette thèse, de nouvelles méthodes pourraient être envisagées pour la validation des données enregistrées par les compteurs et l'identification des compteurs défectueux. Les bornes de validation actuelles ne valident qu'une faisabilité du point de vue physique, et fournissent donc une plage très large de validation. Si dans les années à venir, les estimateurs développés témoignent d'une grande fiabilité, ceux-ci pourraient être utilisés pour valider les consommations futures enregistrées par les compteurs. Dans le cas où un compteur donnerait des

mesures hors des intervalles de confiance des estimateurs, le compteur pourrait être identifié comme défectueux. En revanche, si plusieurs compteurs étaient identifiés comme défectueux en même temps, on pourrait suspecter qu'une évolution de la consommation modélisée a eu lieu, ce qui impliquerait de revoir les estimateurs développés.

Un sujet important non traité dans cette thèse est l'estimation des pertes électriques générées dans l'infrastructure électrique (e.g. les pertes par effet Joule dans la caténaire) lors que l'on consomme de l'énergie. En effet, dans cette thèse, l'objectif principal a traité de l'énergie consommée par les trains au niveau du pantographe. Néanmoins, des quantités d'énergie considérables (jusqu'au 25 % en alimentation 1,5 kV DC) sont perdues dans l'infrastructure électrique et elles sont très dépendantes de la position relative du train et des sous-stations. Une extension de l'étude intégrant les outils développés dans cette thèse pourrait être pertinente.

CONCLUSION ET PERSPECTIVES

Liste des publications

Cette thèse a donné lieu aux publication suivantes :

Articles de revue :

- L.M. ALONSO, L.D. ROUX, L. TAUNAY, A. WATARE, C. SAUDEMONT & B. ROBYNS, « Energy metering data estimation and validation in railways », *IEEE Transactions on Power Delivery*, févr. 2022. Accepté pour publication, doi : 10.1109/TPWRD.2022.3149727

Conférences internationales :

- L.M. ALONSO, A. WATARE, L.D. ROUX, L. TAUNAY, C. SAUDEMONT & B. ROBYNS, « New approach for missing consumption reconstruction and validation of energy-meters data », présenté à *2020 17th International Conference on Railway Engineering Design & Operation (COMPRAIL)*, juil. 2020.
- L.M. ALONSO, A. WATARE, L.D. ROUX, L. TAUNAY, C. SAUDEMONT & B. ROBYNS, « Energy-meter data validation in railway via physical models », présenté à *2020 IEEE Vehicular Power and Propulsion Conference*, nov-dec 2020.

LISTE DES PUBLICATIONS

Bibliographie

- [1] A. Utgaard et D. M. Gulbrandsen, “Eress : A New Era for Railway Energy,” n°. March, p. 6–11, 2020.
- [2] J.-M. Allenbach, *Techniques Ferroviaires*, 2009. [En ligne]. Disponible : <http://www.traction-electrique.ch/>
- [3] J.-M. Allenbach, P. Chapas, M. Comte et R. Kaller, *TRACTION ÉLECTRIQUE*, 2017.
- [4] Direction du Matériel S.N.C.F, “La dynamique du mouvement des trains,” 1976.
- [5] P. Lukaszewicz, “Running resistance - Results and analysis of full-scale tests with passenger and freight trains in Sweden,” *Proceedings of the Institution of Mechanical Engineers, Part F : Journal of Rail and Rapid Transit*, vol. 221, n°. 2, p. 183–193, 2007.
- [6] P. Lukaszewicz, “Energy Consumption and Running Time for Trains. Modelling of running resistance and driver resistance based on full-scale testing,” Thèse de doctorat, 2001.
- [7] B. P. Rochard et F. Schmid, “A review of methods to measure and calculate train resistances,” *Proceedings of the Institution of Mechanical Engineers, Part F : Journal of Rail and Rapid Transit*, vol. 214, n°. 4, p. 185–199, 2000.
- [8] SNCF Réseau, “Données nécessaires au calcul des marches , à la détermination des charges remorquables et des capacités d’accélération,” 2017.
- [9] R. Bosquet, “Modélisation énergétique et identification des trains pour l’écoconception des lignes ferroviaires à grande vitesse,” Thèse de doctorat, 2015.
- [10] Fundación de los Ferrocarriles Españoles, “Metodología de cálculo del consumo de energía de los trenes de viajeros y actuaciones en el diseño del material rodante para su reducción,” Rapport technique, 2010.
- [11] A. Jakubowski, “Constant vs . Variable Efficiency of Electric Drive in Train Run Simulations,” *2019 26th International Workshop on Electric Drives : Improvement in Efficiency of Electric Drives (IWED)*, p. 1–6, 2019.

- [12] G. M. Scheepmaker, R. M. Goverde et L. G. Kroon, “Review of energy-efficient train control and timetabling,” *European Journal of Operational Research*, vol. 257, n^o. 2, p. 355–376, 2017. [En ligne]. Disponible : <http://dx.doi.org/10.1016/j.ejor.2016.09.044>
- [13] A. Fernandez-Rodriguez, “Design of Robust and Energy-Efficient ATO Speed Profiles of Metropolitan Lines Considering Train Load Variations and Delays,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, n^o. 4, p. 2061–2071, 2015.
- [14] N. Ghaviha, M. Bohlin, F. Wallin et E. Dahlquist, “Optimal Control of an EMU Using Dynamic Programming and Tractive Effort as the Control Variable,” *Proceedings of the 56th Conference on Simulation and Modelling (SIMS 56), October, 7-9, 2015, Linköping University, Sweden*, vol. 119, p. 377–382, 2015.
- [15] G. Matsuura et M. Miyatake, “Optimal train speed profiles by dynamic programming with parallel computing and the fine-tuning of mesh,” *WIT Transactions on the Built Environment*, vol. 135, p. 767–777, 2014.
- [16] H. Ko, T. Koseki et M. Miyatake, “Application of dynamic programming to the optimization of the running profile of a train,” *Computers in Railway Six, 2004*, vol. 15, p. 103–112, 2004.
- [17] V.-D. Doan, H. Fujimoto, T. Koseki, T. Yasuda, H. Kishi et T. Fujita, “Iterative Dynamic Programming for Optimal Control Problem with Isoperimetric Constraint and Its Application to Optimal Eco-driving Control of Electric Vehicle,” *IEEEJ Journal of Industry Applications*, vol. 7, n^o. 1, p. 80–92, 2017.
- [18] J. Lesel, G. Claisse, P. Debay et B. Robyns, “Design of daily energy optimal timetables for metro lines using metaheuristics,” *Proceedings of the 18th Mediterranean Electrotechnical Conference : Intelligent and Efficient Technologies and Services for the Citizen, MELECON 2016*, n^o. April, p. 1–6, 2016.
- [19] J. Lésel, “Optimisation de la consommation énergétique d’une ligne de métro automatique prenant en compte les aléas de trafic à l’aide d’outils d’intelligence artificielle,” Thèse de doctorat, 2016.
- [20] M. H. Bigharaz, A. Afshar, A. Suratgar et F. Safaei, *Simultaneous optimization of energy consumption and train performances in electric railway systems*. IFAC, 2014, vol. 19, n^o. 3. [En ligne]. Disponible : <http://dx.doi.org/10.3182/20140824-6-ZA-1003.02368>
- [21] V. De Martinis et M. Gallo, “Models and Methods to Optimise Train Speed Profiles with and without Energy Recovery Systems : A Suburban Test Case,” *Procedia - Social and Behavioral Sciences*, vol. 87, p. 222–233, 2013. [En ligne]. Disponible : <http://dx.doi.org/10.1016/j.sbspro.2013.10.606>

- [22] I. Asnis, A. Dmitruk et N. Osmolovskii, "Solution of the problem of the energetically optimal control of the motion of a train by the maximum principle," *USSR Computational Mathematics and Mathematical Physics*, vol. 25, n^o. 6, p. 37–44, jan 1985.
- [23] P. Howlett, I. Milroy et P. Pudney, "Energy-efficient train control," *Control Engineering Practice*, vol. 2, n^o. 2, p. 193–200, apr 1994.
- [24] P. Pudney et P. Howlett, "Optimal driving strategies for a train journey with speed limits," *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, vol. 36, n^o. 01, p. 38, jul 1994.
- [25] E. Khmelnitsky, "On an optimal control problem of train operation," *IEEE Transactions on Automatic Control*, vol. 45, n^o. 7, p. 1257–1266, jul 2000.
- [26] LAAS, "Méthodes directes et méthodes indirectes." [En ligne]. Disponible : https://homepages.laas.fr/arzelier/drupal/sites/homepages.laas.fr.arzelier/files/u117/C3_{_}CO_{_}UPS_{_}0.pdf
- [27] G. M. Scheepmaker, P. Pudney, A. R. Albrecht, R. M. Goverde et P. G. Howlett, "Optimal Running Time Supplement Distribution For Energy-Efficient Train Control," dans *8th International Conference on Railway Operations Modelling and Analysis - RailNorrköping 2019*, 2019.
- [28] P. Wang et R. M. Goverde, "Multiple-phase train trajectory optimization with signalling and operational constraints," *Transportation Research Part C : Emerging Technologies*, vol. 69, p. 255–275, 2016. [En ligne]. Disponible : <http://dx.doi.org/10.1016/j.trc.2016.06.008>
- [29] C. Andre-Philippe, F. Didier, M. Pascal, G. Younes, P. Gwenael et H. Patrick, "Energy Efficiency for rolling stock, from measuring to savings," *2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles and International Transportation Electrification Conference, ESARS-ITEC 2018*, 2019.
- [30] V. Kantikoon et V. Kinnares, "The estimation of electrical energy consumption in abnormal automatic meter reading system using multiple linear regression," *2013 International Conference on Electrical Machines and Systems, ICEMS 2013*, p. 826–830, 2013.
- [31] Z. Su et L. Dong, "Prediction of Tram Energy Consumption Based on Modified Regression Model," *Proceedings - 2019 2nd International Conference on Safety Produce Informatization, IICSPI 2019*, p. 188–193, 2019.
- [32] O. A. Olanrewaju, "Predicting Industrial Sector's Energy Consumption : Application of Support Vector Machine," *IEEE International Conference on Industrial Engineering and Engineering Management*, p. 1597–1600, 2019.

- [33] A. Amara Korba et N. El Islem Karabadji, “Smart Grid Energy Fraud Detection Using SVM,” *Proceedings - ICNAS 2019 : 4th International Conference on Networking and Advanced Systems*, p. 1–6, 2019.
- [34] A. Parlak, Y. Islamoglu, H. Yasar et A. Egrisogut, “Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a Diesel engine,” *Applied Thermal Engineering*, vol. 26, n^o. 8-9, p. 824–828, 2006.
- [35] A. A. Komyakov, M. M. Nikiforov, V. V. Erbes, V. T. Cheremisin et V. I. Ivanchenko, “Construction of electricity consumption mathematical models on railway transport used artificial neural network and fuzzy neural network,” *EEEIC 2016 - International Conference on Environment and Electrical Engineering*, vol. 1, n^o. 1, p. 1–4, 2016.
- [36] A. Lenail, “NN SVG.” [En ligne]. Disponible : <https://alexlenail.me/NN-SVG/index.html> (Visité le 25 octobre 2021).
- [37] “« PLANÈTE Voyages » : pour une mobilité plus écologique| SNCF.” [En ligne]. Disponible : <https://www.sncf.com/fr/offres-voyageurs/voyager-en-train/planete-voyages-mobilite-ecologique> (Visité le 16 septembre 2021).
- [38] “Lorry Rail - L’autoroute ferroviaire Le Boulou - Bettembourg - Services.” [En ligne]. Disponible : <http://www.lorry-rail.com/services/> (Visité le 24 février 2017).
- [39] “Sustainable Energy | About NS | NS.” [En ligne]. Disponible : <http://www.ns.nl/en/about-ns/energy/sustainable-energy.html> (Visité le 8 mars 2017).
- [40] Navigant Consulting, “The Use of Renewable Energy Sources to Provide Power to California’s High Speed Rail,” Rapport technique, 2008.
- [41] P.-V. Buzila, “Gestion énergétique optimale des Installations Fixes de Traction Electrique Ferroviaire Hybrides,” Thèse de doctorat, 2015.
- [42] P. Pankovits, D. Abbes, C. Saudemont, S. Brisset, J. Pouget et B. Robyns, “Multi-criteria fuzzy-logic optimized supervision for hybrid railway power substations,” *Mathematics and Computers in Simulation*, vol. 130, p. 236–250, 2016. [En ligne]. Disponible : <http://dx.doi.org/10.1016/j.matcom.2016.05.002>
- [43] J. B. Robyns, C. Saudemont, D. Hissel, X. Roboam, B. Sareni, *Electrical Energy Storage in Transportation Systems*, ISTE-Wiley, édit., 2016.
- [44] K. Almaksour, Y. Krim, N. Kouassi, N. Navarro, B. François, T. Letrouvé, C. Saudemont, L. Taunay et B. Robyns, “Comparison of dynamic models for a DC railway electrical network

- including an AC/DC bi-directional power station,” *Mathematics and Computers in Simulation*, vol. 184, p. 244–266, 2021. [En ligne]. Disponible : <https://doi.org/10.1016/j.matcom.2020.05.027>
- [45] D. Ramsey, T. Letrouve, A. Bouscayrol et P. Delarue, “Quantification of the recoverable braking energy in a reversible railway DC traction power substation,” *2019 21st European Conference on Power Electronics and Applications, EPE 2019 ECCE Europe*, p. 1–10, 2019.
- [46] T. Letrouvé et G. Gazeignies, “Feasibility study of trains involvement in electric demand-response,” dans *Transport Research Arena 2018*, 2018.
- [47] T. Letrouvé, “Le smart grid ferroviaire : quels degrés de flexibilité dans le matériel roulant ferroviaire ?” *Revue de l’Electricité et de l’Electronique*, vol. 5, p. 1–110, 2020.
- [48] A. Garcia Alvarez et M. d. P. Martin Cañizares, “Metodología de cálculo del consumo de energía de los trenes de viajeros y actuaciones en el diseño del material rodante para su reducción,” vol. 1, p. 1–58, 2010.

BIBLIOGRAPHIE

Annexe 1

Modélisation de la force en courbe (F_c)

Dans la littérature les formules suivantes sont mentionnées pour modéliser la force en courbe :

$$r_c[kN/t] = 0.01 * \frac{k}{R_c} \quad (A1.1)$$

Avec k un paramètre qui varie de 500 à 1200 selon les différentes caractéristiques du train (e.g. distance entre essieux dans les bogies) avec une valeur moyenne de 800 et R_c le rayon de courbure en mètres.

$$r_c[kN/t] = \frac{1.6 * a + 1.62}{R_c} \quad (A1.2)$$

Avec a la distance entre les essieux d'un bogie en mètres et R_c le rayon de courbure en mètres.

$$r_c[daN/t] = \frac{600}{R_c} * (1 - 0.2K_{rlibres} + 0.03 * (D_{ejes} - 3) - 0.4 * k_{rradiales}) \quad (A1.3)$$

Avec R_c le rayon de courbure en mètres, $K_{rlibres}$ (1 si roues libres (type Talgo) 0 si roues sur un axe), D_{ejes} distance entre essieux et $k_{rradiales}$ (0 si roues orientables (type Talgo) 0 sinon)[48].

Modélisation des coefficients A, B et C de la force de résistance à l'avancement (F_{res})

Dans la littérature, plusieurs propositions pour estimer la valeur des coefficients A, B et C sont disponibles :

- **Armstrong & Swift coefficients** [7] :

$$\begin{aligned}
 A &= 6.4 * (m_{remorques}) + 8.0 * (m_{motrices}) \\
 B &= 0.18 * ((m_{train}) + (n_{remorques})) + 0.005 * (n_{motrices}) * (p_{tot}) \\
 C &= 0.6125 * (c_{trainage}) * (s_{frontale}) + 0.00197 * (p_{frontale} * l) \\
 &\quad + 0.0021 * (p_{frontale} * d_{voitures} * (n_{motrices} + n_{remorques} - 1)) \\
 &\quad + 0.2061 * (c_{bogie}) * (n_{bogie}) + 0.2566 * (n_{panto})
 \end{aligned} \tag{A1.4}$$

Avec $m_{remorques}$ la masse des remorques en tonnes, $m_{motrices}$ la masse des machines en tonnes, m_{train} la masse totale du train (machines plus remorques) en tonnes, $n_{remorques}$ le nombre de remorques, $n_{motrices}$ le nombre de machines, p_{tot} puissance totale des machines en kW, $c_{trainage}$ le coefficient de trainage de la tête/queue, $s_{frontale}$ la surface frontale en m², $p_{frontale}$ le périmètre frontale en mètres, $d_{voitures}$ l'espace entre voitures en mètres, c_{bogie} le coefficient de trainage du bogie, n_{bogie} le nombre de bogies et n_{panto} le nombre de pantographes dépliés du train.

- **France (SNCF)** (force en kN, vitesse en m/s)[4]

$$\begin{aligned}
 A &= 0.00001 * \lambda * M * \sqrt{\frac{10000}{m}} \\
 B &= 3.6 * 10^{-7} * M \\
 C &= 0.1296 * (k_1 * S + k_2 * p * L)
 \end{aligned} \tag{A1.5}$$

Avec M la masse totale du train (kg), m la charge par essieu (kg) et λ paramètre entre 0.9 (TGV) et 1.5 (FRET). Avec p le périmètre partiel (autour 10 m normalement), L la longueur du train, k_1 un paramètre dépendant du profil aérodynamique des faces avant/arrière du train (entre $20 * 10^{-4}$ pour le matériel conventionnel et $9 * 10^{-4}$ pour les TGV) et k_2 un paramètre dépendant de la condition de la surface (entre $30 * 10^{-6}$ pour le matériel conventionnel et $20 * 10^{-6}$ pour les TGV).

La SNCF avait par ailleurs proposé des formules simplifiées pour divers types de matériels, ces formules calculent r (en N/kg) la résistance à l'avancement :

- Voitures sur bogies :

$$r = 0.015 + 0.000036 * V^2 \quad (A1.6)$$

- Voitures UIC :

$$r = 0.0125 + 0.000021 * V^2 \quad (A1.7)$$

- Voitures de 2 ou 3 essieux et FRET express :

$$r = 0.015 + 0.000054 * V^2 \quad (A1.8)$$

- FRET avec 10t de charge par essieu :

$$r = 0.015 + 0.000081 * V^2 \quad (A1.9)$$

- FRET avec 18t de charge par essieu :

$$r = 0.015 + 0.0000324 * V^2 \quad (A1.10)$$

- Machines électriques ou diesel, R (kN) :

$$r = 0.0065 * L + 0.13 * n + 0.0036 * V + 0.003888 * V^2 \quad (A1.11)$$

Avec L le poids de la machine en tonnes et n le nombre d'essieux.

Pour les EMU (Unité Electrique Multiple) :

$$r = \left(0.013 * \sqrt{\frac{10}{m}} \right) + 0.00036 * V * P + 0.1296 * C * V^2 \quad (A1.12)$$

$$C = 0.0035 * S + 0.0041 * \frac{p * L}{100} + 0.002 * N$$

Avec P le poids total de l'EMU, m la charge par essieu et N le nombre de pantographes dépliés.

Finalement, dans le cas d'un TGV

$$r = 25 + 1.188 * V + 0.0703728 * V^2 \quad (A1.13)$$

-**Allemagne (DB)** (force en kN, vitesse en m/s) pour des trains classiques (locomotives + voitures) ou trains FRET :

$$r = 0.01 * \left((1.5 - 2.5) + (0.007 + 1/m)((3.6 * V)/10)^2 \right) \quad (A1.14)$$

Avec :

$m = 40$ pour voitures de 4 essieux ou wagons FRET

m = 30 pour voitures de 3 ou 2 essieux
m = 20 pour FRET express
m = 10 pour wagons FRET sans charge

Une autre proposition pour des trains passagers :

$$r = 1.9 + 3.6 * b * V + 0.0048 * \frac{1}{Q} * (n + 2.7) * A * (3.6 * V + 15)^2 \quad (\text{A1.15})$$

Avec A l'aire équivalente ($1.45m^2$ pour des TGV, $1.55 m^2$ pour des trains rapides et $1.15m^2$ pour des trains avec 2 ou 3 essieux). $b = 0.0025$ pour des wagons à 4 essieux, 0.004 3 essieux et 0.007 2 essieux. Q le poids du train (t).

-Espagne (RENFE) (force en daN, vitesse en km/h)[48]

$$A = [(M + n_v * 0.08) * 0.25] + 7 * N_e \quad (\text{A1.16})$$

Avec M la masse à vide du train, n_v le nombre de voyageurs et N_e le nombre d'essieux.

$$B = \left(\frac{P_{train}}{243} + n_p * 0.14166 \right) * 0.034 \quad (\text{A1.17})$$

Avec P_{train} la puissance totale du train en kW et n_p le nombre de places disponibles dans le train.

$$C = (W_{ext} * H_{ext_{max}}) * 0.00096 + [(2 * H_{ext_{med}} + W_{ext}) * l_{tren}] * 0.000021 + 0.0002 * P_{ant} \quad (\text{A1.18})$$

Avec W_{ext} la largeur extérieure moyenne, $H_{ext_{max}}$ la hauteur extérieure maximale, $H_{ext_{med}}$ la hauteur extérieure moyenne, l_{tren} la longueur du train et P_{ant} le nombre de pantographes déployés.

Annexe 2

Cas d'étude FRET n°1 Avignon-Valenton

Nom de la sous-station	Numéro de la sous-station
AISEREY	1
AMPUIS	2
BAULME	3
BEAUNE	4
BEDARRIDES	5
BELLEVILLE	6
BENY	7
BETTANT	8
BRUNOY	9
CESSON	10
CEZY	11
CHAGNY	12
CHALON	13
CHAMP ROLLAND	14
CHAMP DE LAYGUE	15
CHAMPFLEURY	16
CHAMPIGNY	17
CHANOZ	18
CHASSE	19
CHATEAUNEUF	20
CLONAS	21
COMBEAUX	22
COMBS	23
CORGOLOIN	24
CRECHES	25
CREMADE	26
CROSNE	27
FLACE	28
FLEURVILLE	29
FLOGNY	30

FONTAINES	31
FRONTENAUD	32
GEVREY	33
GISSEY	34
GUILHERAND	35
LA COMBETTE	36
LA GRANGE	37
LA GUILLOTIERE	38
LA PALUD	39
LA RONCIERE	40
LA SAULAIE	41
LAROCHE	42
LE PAIREUX	43
LE PONTET	44
LE POUZIN	45
LE VILLARS	46
LES CADALLES	47
LES GRANDS VIOLETS	48
LES OLIVETTES	49
LES TARTERETS	50
LIEUSAIN	51
LIMONY	52
LORIOLE	53
MELUN	54
MERVANS	55
MEURSAULT	56
MEYSSE	57
MONDRAGON	58
MONTEREAU	59
MONTLUEL	60
NAVILLY	61
NUITS	62
OZON	63
PACY	64
PERRIGNY	65
PETIT BERU	66
PIOLENC	67
POLLIAT	68
PORTES	69
QUINCIEUX	70
REBOUL	71
RIS ORANGIS	72
ROMANECHE	73

SAMOIS	74
SAMOREAU	75
SAULCE	76
SEIGNY	77
SENNECEY	78
SENOZAN	79
SERVES	80
SIBELIN	81
ST AMOUR	82
ST FLORENTIN	83
ST GENIES	84
ST GEORGES	85
ST MAMMES	86
ST MARTIN DU TERTRE	87
ST NAZAIRE	88
ST USUGE	89
ST VALLIER	90
STE BARBE	91
TAIN	92
TOURNON	93
VARENNES	94
VAUGRIS	95
VELARS	96
VERREY	97
VILLEFRANCHE	98
VILLENEUVE SUR YONNE	99
VILLERS	100
VIVIERS	101
VONNAS	102
VOSNE	103

TABLEAU A2.3 – Apport des sources d'information les plus importantes pour les estimateurs les plus performants pour le cas d'étude Avignon-Valenton

Avignon-Valenton SVM 12,07 %	Avignon-Valenton avec sous-stations RML 11,68 %	Avignon-Valenton avec sous-stations et masse RML 11,48 %
Heure départ 46,91 %	Heure départ 40,72 %	Heure départ 38,80 %
Durée trajet 37,04 %	Durée trajet 22,41 %	Durée trajet 21,26 %
Mois 13,82 %	STA 65 10,27 %	STA 65 10,06 %
Jour année 2,23 %	STA 33 10,10 %	STA 33 9,54 %
WE/Ferie Bruit	STA 11 3,29 %	TBR 5,43 %
Direction Bruit	STA 2 2,21 %	STA 42 3,51 %
Température à Orly Bruit	STA 22 1,35 %	STA 50 1,64 %
Température à Montelimar Bruit	Mois 1,16 %	TBC 1,12 %
UM Bruit	STA 101 0,81 %	STA 101 1,03 %
Jour semaine Bruit	STA 41 0,77 %	STA 81 0,93 %
	STA 38 0,77 %	STA 74 0,88 %
	STA 28 0,72 %	STA 14 0,76 %
	STA 74 0,72 %	STA 38 0,74 %
	STA 81 0,68 %	STA 89 0,67 %
	STA 71 0,67 %	STA 28 0,65 %
	STA 50 0,65 %	STA 72 0,53 %
	STA 58 0,51 %	STA 92 0,52 %
	STA 39 0,51 %	STA 39 0,49 %
	STA 46 0,39 %	STA 58 0,48 %

	STA 89 0,39 %	STA 41 0,38 %
	STA 44 0,35 %	Température à Orly 0,31 %
	STA 14 0,26 %	STA 45 0,27 %
	STA 9 0,24 %	STA 52 0,27 %
	WE/Ferie 0,22 %	STA 90 0,25 %
	STA 90 0,15 %	STA 94 0,17 %
	STA 83 0,11 %	STA 49 0,15 %
	STA 72 0,10 %	STA 18 0,13 %
	STA 16 0,10 %	STA 9 0,01 %
	STA 86 0,08 %	STA 40 Bruit
	STA 82 0,07 %	STA 46 Bruit
	STA 31 0,06 %	STA 44 Bruit
	STA 30 0,06 %	STA 27 Bruit
	STA 63 0,05 %	STA 76 Bruit
	STA 45 0,01 %	STA 103 Bruit
	STA 42 Bruit	STA 3 Bruit
	STA 76 Bruit	STA 29 Bruit
	STA 21 Bruit	STA 16 Bruit
	STA 20 Bruit	STA 11 Bruit
	STA 40 Bruit	STA 25 Bruit
	STA 13 Bruit	STA 5 Bruit
	STA 95 Bruit	WE/Ferie Bruit

	Température à Orly Bruit	STA 34 Bruit
	STA 94 Bruit	STA 69 Bruit
	Température à Montelimar Bruit	STA 60 Bruit
	STA 97 Bruit	STA 32 Bruit
	UM Bruit	STA 98 Bruit
	STA 61 Bruit	STA 8 Bruit
	STA 75 Bruit	STA 26 Bruit
	STA 66 Bruit	STA 19 Bruit
	STA 8 Bruit	STA 59 Bruit
	STA 26 Bruit	STA 93 Bruit
	STA 3 Bruit	STA 13 Bruit
	Direction Bruit	STA 21 Bruit
	STA 69 Bruit	STA 2 Bruit
	STA 78 Bruit	STA 96 Bruit
	STA 67 Bruit	UM Bruit
	STA 80 Bruit	STA 95 Bruit
	STA 48 Bruit	STA 4 Bruit
	STA 18 Bruit	STA 17 Bruit
	STA 43 Bruit	STA 31 Bruit
	STA 70 Bruit	STA 82 Bruit
	STA 64 Bruit	STA 67 Bruit
	STA 17 Bruit	STA 84 Bruit

	Jour semaine Bruit	STA 66 Bruit
	STA 49 Bruit	STA 70 Bruit
	STA 52 Bruit	STA 64 Bruit
	STA 12 Bruit	STA 48 Bruit
	STA 37 Bruit	STA 43 Bruit
	STA 68 Bruit	STA 12 Bruit
	STA 56 Bruit	STA 87 Bruit
	STA 60 Bruit	STA 30 Bruit
	STA 87 Bruit	STA 61 Bruit
	STA 36 Bruit	STA 35 Bruit
	Jour année Bruit	STA 51 Bruit
	STA 88 Bruit	Mois Bruit
	STA 98 Bruit	STA 80 Bruit
	STA 24 Bruit	STA 7 Bruit
	STA 55 Bruit	STA 71 Bruit
	STA 51 Bruit	STA 53 Bruit
	STA 96 Bruit	STA 20 Bruit
	STA 15 Bruit	STA 54 Bruit
	STA 7 Bruit	STA 86 Bruit
	STA 84 Bruit	STA 57 Bruit
	STA 4 Bruit	STA 97 Bruit
	STA 5 Bruit	Jour semaine Bruit

	STA 93 Bruit	Température à Montelimar Bruit
	STA 34 Bruit	STA 15 Bruit
	STA 32 Bruit	STA 22 Bruit
	STA 99 Bruit	STA 78 Bruit
	STA 25 Bruit	STA 91 Bruit
	STA 73 Bruit	STA 56 Bruit
	STA 92 Bruit	STA 75 Bruit
	STA 62 Bruit	STA 79 Bruit
	STA 77 Bruit	Jour année Bruit
	STA 27 Bruit	STA 62 Bruit
	STA 59 Bruit	Direction Bruit
	STA 1 Bruit	STA 47 Bruit
	STA 100 Bruit	STA 83 Bruit
	STA 91 Bruit	STA 55 Bruit
	STA 47 Bruit	STA 99 Bruit
	STA 35 Bruit	STA 24 Bruit
	STA 103 Bruit	STA 77 Bruit
	STA 53 Bruit	STA 68 Bruit
	STA 57 Bruit	STA 63 Bruit
	STA 29 Bruit	STA 73 Bruit
	STA 54 Bruit	STA 36 Bruit
	STA 19 Bruit	STA 1 Bruit

	STA 102 Bruit	STA 37 Bruit
	STA 79 Bruit	STA 88 Bruit
	STA 85 Bruit	STA 100 Bruit
	STA 6 Bruit	STA 102 Bruit
		STA 6 Bruit
		STA 85 Bruit

Cas d'étude FRET n°2 Le Boulou - Zoufftgen

Nom de la sous-station	Numéro de la sous-station
AISEREY	1
AMPUIS	2
BEAUNE	3
BEDARRIDES	4
BELLEVILLE	5
BENY	6
BETTANT	7
BEZIERS	8
BISCARRAT	9
CASTELLAS	10
CASTELNAU	11
CHAGNY	12
CHALON	13
CHAMP DE LAYGE	14
CHAMP ROLLAND	15
CHAMP SALMON	16
CHAMPFLEURY	17
CHAMPS REGNAUDS	18
CHANOZ	19
CHASSE	20
CHATEAUNEUF	21
CLONAS	22
COMBEAUX	23
CORGOLOIN	24
CRECHES	25
CREMADE	26
FIRMINELLES	27
FITOU	28
FLACE	29
FLEURVILLE	30
FONTAINES	31
FRONTENAUD	32
GEVREY	33
GONCOURT	34
GRISE COTE	35
GUILHERAND	36
LA CASTELLE	37
LA COMBETTE	38
LA GRANGE	39
LA GUILLOTIERE	40

LA NADIERE	41
LA PALUD	42
LA RONCIERE	43
LA SAULAIE	44
LE BOULOU	45
LE PAIREUX	46
LE PONTET	47
LE POUZIN	48
LE RELE	49
LE VILLARS	50
LEDENON	51
LEROUVILLE	52
LES BRUYERES	53
LES CADALLES	54
LES GRANDS VIOLETS	55
LES OLIVETTES	56
LIMONY	57
LORIOI	58
LUNEL VIEL	59
MASSANE	60
MERVANS	61
MEURSAULT	62
MEYSSE	63
MIREVAL	64
MONDRAGON	65
MONTLUEL	66
NARBONNE	67
NAVILLY	68
NIMES	69
NISSAN	70
OZON	71
PANISSO	72
PERPIGNAN	73
PERRIGNY	74
PIOLENC	75
POLLIAT	76
PORTES	77
QUINCIEUX	78
REBOUL	79
ROMANECHÉ	80
SAULCE	81
SENNECEY	82
SENOZAN	83

SERVES	84
SETE	85
SIBELIN	86
ST AMOUR	87
ST GENIES	88
ST GEORGES	89
ST NAZAIRE	90
ST USUGE	91
ST VALLIER	92
TAIN	93
THEZIER	94
THIONVILLE	95
TOURNON	96
VARENNES	97
VAUGRIS	98
VESTRIC	99
VIAS	100
VILLEFRANCHE	101
VILLEGUSIEN	102
VIVIERS	103
VONNAS	104
VOSNE	105

TABLEAU A2.5 – Apport des sources d'information les plus importantes pour les estimateurs les plus performants pour le cas d'étude Le Boulou-Zoufftgen

Le Boulou-Zoufftgen SVM 26,72%	Le Boulou-Zoufftgen avec sous-stations RLM 26,29%	Le Boulou-Zoufftgen avec sous-stations et masse RLM 26,39%
UM 93.64 %	UM 65.75 %	UM 66.89 %
Direction 5.10 %	STA 69 10.72 %	STA 69 10.53 %
Heure départ 1.27 %	STA 88 4.35 %	STA 88 4.03 %
Jour semaine 0.78 %	STA 56 2.84 %	STA 94 2.50 %
Mois Bruit	STA 89 2.24 %	STA 89 2.14 %
Durée trajet Bruit	STA 81 2.12 %	STA 56 2.01 %
WE/Ferie Bruit	STA 74 2.01 %	STA 5 1.99 %
Température à Perpignan Bruit	STA 94 1.95 %	STA 15 1.73 %
Température à Nancy Bruit	STA 45 1.54 %	STA 81 1.63 %
	STA 5 1.52 %	Heure départ 1.48 %
	STA 64 1.40 %	STA 28 1.44 %
	STA 15 1.31 %	STA 45 1.36 %
	Heure départ 1.19 %	STA 50 1.15 %
	STA 53 0.73 %	STA 53 0.85 %
	STA 79 0.62 %	STA 99 0.84 %
	STA 93 0.46 %	STA 64 0.82 %
	STA 54 0.39 %	STA 79 0.65 %
	STA 99 0.38 %	TBC 0.59 %
	STA 28 0.37 %	STA 26 0.59 %

	STA 71 0.34 %	STA 71 0.55 %
	STA 67 0.33 %	STA 70 0.42 %
	STA 29 0.24 %	WE/Ferie 0.36 %
	Température à Nancy 0.23 %	TBR 0.19 %
	STA 18 0.22 %	STA 41 0.15 %
	STA 7 0.17 %	STA 6 0.02 %
	STA 90 0.16 %	Température à Perpignan 0.01 %
	STA 96 0.09 %	STA 83 Bruit
	STA 23 0.09 %	Durée trajet Bruit
	WE/Ferie 0.07 %	STA 27 Bruit
	STA 70 Bruit	STA 14 Bruit
	Température à Perpignan Bruit	STA 30 Bruit
	STA 78 Bruit	STA 101 Bruit
	STA 24 Bruit	STA 78 Bruit
	STA 47 Bruit	STA 84 Bruit
	STA 20 Bruit	STA 93 Bruit
	STA 40 Bruit	STA 57 Bruit
	STA 6 Bruit	STA 65 Bruit
	STA 41 Bruit	STA 82 Bruit
	STA 50 Bruit	STA 97 Bruit
	STA 57 Bruit	STA 4 Bruit
	STA 3 Bruit	STA 20 Bruit

	STA 62 Bruit	STA 25 Bruit
	STA 75 Bruit	STA 47 Bruit
	STA 55 Bruit	STA 74 Bruit
	STA 38 Bruit	STA 29 Bruit
	Mois Bruit	STA 62 Bruit
	STA 83 Bruit	STA 32 Bruit
	STA 72 Bruit	STA 48 Bruit
	STA 84 Bruit	STA 18 Bruit
	STA 42 Bruit	STA 7 Bruit
	STA 14 Bruit	STA 34 Bruit
	STA 26 Bruit	STA 24 Bruit
	STA 77 Bruit	STA 33 Bruit
	STA 91 Bruit	STA 38 Bruit
	Direction Bruit	STA 12 Bruit
	STA 25 Bruit	STA 2 Bruit
	STA 82 Bruit	STA 98 Bruit
	STA 27 Bruit	Jour année Bruit
	STA 30 Bruit	STA 87 Bruit
	Durée trajet Bruit	STA 86 Bruit
	STA 73 Bruit	Mois Bruit
	STA 87 Bruit	STA 67 Bruit
	STA 1 Bruit	STA 100 Bruit

	STA 8 Bruit	STA 73 Bruit
	STA 51 Bruit	STA 63 Bruit
	STA 10 Bruit	STA 42 Bruit
	STA 16 Bruit	STA 21 Bruit
	STA 39 Bruit	Jour semaine Bruit
	STA 2 Bruit	STA 11 Bruit
	STA 98 Bruit	STA 40 Bruit
	STA 11 Bruit	STA 95 Bruit
	STA 4 Bruit	STA 10 Bruit
	STA 21 Bruit	STA 8 Bruit
	STA 58 Bruit	STA 55 Bruit
	STA 31 Bruit	STA 72 Bruit
	STA 95 Bruit	STA 23 Bruit
	STA 19 Bruit	STA 102 Bruit
	STA 46 Bruit	STA 60 Bruit
	STA 97 Bruit	STA 1 Bruit
	STA 63 Bruit	STA 46 Bruit
	STA 32 Bruit	Direction Bruit
	Jour semaine Bruit	STA 44 Bruit
	STA 60 Bruit	STA 39 Bruit
	STA 65 Bruit	STA 54 Bruit
	STA 22 Bruit	STA 19 Bruit

	STA 48 Bruit	STA 16 Bruit
	Jour année Bruit	STA 43 Bruit
	STA 52 Bruit	STA 77 Bruit
	STA 68 Bruit	STA 96 Bruit
	STA 59 Bruit	STA 22 Bruit
	STA 13 Bruit	STA 85 Bruit
	STA 17 Bruit	STA 68 Bruit
	STA 44 Bruit	STA 91 Bruit
	STA 80 Bruit	STA 9 Bruit
	STA 9 Bruit	STA 66 Bruit
	STA 33 Bruit	Température à Nancy Bruit
	STA 43 Bruit	STA 58 Bruit
	STA 92 Bruit	STA 36 Bruit
	STA 86 Bruit	STA 51 Bruit
	STA 34 Bruit	STA 92 Bruit
	STA 12 Bruit	STA 75 Bruit
	STA 36 Bruit	STA 80 Bruit
	STA 66 Bruit	STA 103 Bruit
	STA 61 Bruit	STA 105 Bruit
	STA 85 Bruit	STA 59 Bruit
		STA 3 Bruit
		STA 52 Bruit

ANNEXE 2

		STA 13 Bruit
		STA 31 Bruit
		STA 17 Bruit
		STA 90 Bruit
		STA 61 Bruit

Cas d'étude FRET n°3 Dunkerque - Ebange

Nom de la sous-station	Numéro de la sous-station
BUSIGNY	1
CHAMP SALMON	2
CHAUVENCY	3
COUDEKERQUE	4
DELIVRANCE	5
FOURMIES	6
HAUMONT	7
JOUDREVILLE	8
LA FLAMENGRIE	9
LES TERRES NOIRES	10
MAING	11
MOHON	12
RUMINGHEM	13
THIONVILLE	14
VERQUIGNEUL	15

TABLEAU A2.6 – Apport des sources d'information les plus importantes pour les estimateurs les plus performants pour le cas d'étude Dunkerque-Ebange

Dunkerque-Ebange SVM 14,15%	Dunkerque-Ebange avec sous-stations SVM 11,72%	Dunkerque-Ebange avec sous-stations et masse SVM 11,77%
Durée trajet 46.34 %	STA 15 27.76 %	STA 14 24.27 %
Heure 32.48 %	STA 14 13.95 %	Température à Lille 14.41 %
Température à Lille 6.69 %	STA 11 12.77 %	STA 15 14.01 %
Jour semaine 4.42 %	STA 8 12.54 %	STA 11 12.86 %
Mois 3.90 %	Mois 12.00 %	STA 1 12.02 %
Température à Nancy 3.46 %	Heure 7.20 %	Jour année 10.68 %
Jour année 2.71 %	Température à Nancy 6.72 %	STA 9 7.04 %
Direction Bruit	STA 1 5.93 %	STA 3 2.63 %
UM Bruit	Jour année 1.13 %	Mois 2.05 %

WE/Ferie Bruit	Direction 0.12 %	TBC 0.50 %
	STA 10 Bruit	STA 10 0.07 %
	STA 5 Bruit	STA 7 0.03 %
	STA 9 Bruit	Direction Bruit
	STA 12 Bruit	STA 6 Bruit
	STA 3 Bruit	STA 5 Bruit
	STA 6 Bruit	Heure Bruit
	STA 7 Bruit	STA 4 Bruit
	STA 4 Bruit	TBR Bruit
	UM Bruit	STA 8 Bruit
	Température à Lille Bruit	Température à Nancy Bruit
	WE/Ferie Bruit	STA 12 Bruit
	Durée trajet Bruit	Durée trajet Bruit
	STA 13 Bruit	WE/Ferie Bruit
	STA 2 Bruit	UM Bruit
	Jour semaine Bruit	Jour semaine Bruit
		STA 13 Bruit
		STA 2 Bruit

Résumé : Les compteurs installés dans les trains pour le comptage d'énergie dans le contexte de l'ouverture à la concurrence du marché électrique et ferroviaire présentent divers dysfonctionnements. Pour répondre à la problématique de la fiabilité des compteurs, dans cette thèse, des bornes de validation des consommations relevées par les compteurs basés sur des modèles physiques ont été développées. Ces bornes permettent d'augmenter la fiabilité des données de comptage en écartant les mesures aberrantes du point de vue physique. Afin d'améliorer la disponibilité des données de comptage, des estimateurs de la consommation sont développés dans cette thèse, en utilisant une base de données composée par des mesures de compteurs, réalisées sur une ou plusieurs années, validées par les bornes définies précédemment. L'impact de l'évolution du réseau ferroviaire vers un smart grid ferroviaire sur la consommation ferroviaire et le comptage de l'énergie est abordé en fin de thèse.

Mots clés : Comptage énergétique, ferroviaire, validation de la consommation, estimation de la consommation.

Abstract : The energy-meters installed in trains in the context of the opening to competition of the electricity and railway market present various malfunctions. In order to address the issue of meter reliability, in this thesis, validation bounds for energy-meter readings based on physical models have been developed. These bounds allow to increase the reliability of the metering data by discarding physically outlier measurements. To improve the availability of metering data, estimators of the consumption are developed in this thesis, using a database composed of energy-meter measurements, carried out over one or more years and validated by the previously defined bounds. The impact of the evolution of the railway network towards a railway smart grid on railway consumption and energy metering is discussed at the end of the thesis.

Keywords : Energy metering, railways, consumption validation, consumption estimation.